Eye tracking in the cockpit: a review of the relationships between eye movements and the aviator's cognitive state

Mackenzie G. Glaholt
DRDC – Toronto Research Centre

Defence Research and Development Canada

Scientific Report
DRDC-RDDC-2014-R153
December 2014

Eye tracking in the cockpit: a review of the relationships between eye movements and the aviator's cognitive state

Mackenzie G. Glaholt
DRDC – Toronto Research Centre

Defence Research and Development Canada

Scientific Report
DRDC-RDDC-2014-R153
December 2014

Abstract

The eye movements of aircrew during flight have been a topic of interest to military and civilian researchers for over 60 years. Studies in flight simulators and real aircraft have used eye movements as a window onto operators' processing of information from cockpit instruments and displays. This body of research has demonstrated strong links between eye movements and pilot expertise, workload, and situational awareness. Other applications of eye tracking in the cockpit have also been proposed, such as usability analysis, instructor-feedback, and aircraft input. The present report reviews this body of research to date and provides recommendations for future research and applications.

Significance to defence and security

The Royal Canadian Air Force (RCAF) is committed to increasing operational effectiveness of aircrew. This may be achieved through the application of novel technologies that enhance selection, training, or that directly enhance operational performance. This review demonstrates that eye tracking applications have the potential to enhance both training and operational performance in the context of military flight. This report was requested by The Director Air Personnel Strategy (D Air Pers Strat), and is the final deliverable for the Defence Research and Development Canada (DRDC) Project 03rm ("Eye Tracking Technology in Next Generation Pilot and Aircrew Training").

Résumé

Les mouvements oculaires des équipages en vol suscitent l'intérêt des chercheurs militaires et civils depuis plus de 60 ans. Au cours d'études menées en simulateurs de vol et à bord d'aéronefs, ils ont servi à analyser la façon dont les pilotes traitent l'information fournie par les instruments et les dispositifs d'affichage de leur habitacle. Ces travaux ont montré qu'il existe des liens étroits entre le mouvement des yeux des pilotes et leur savoir-faire, leur charge de travail et leur connaissance de la situation. D'autres applications de suivi des mouvements oculaires ont été proposées, celles-ci portant notamment sur l'analyse de la facilité d'utilisation, sur la rétroaction d'instructeurs et sur les données fournies par les aéronefs. Le présent rapport traite de la recherche effectuée jusqu'à maintenant à ce chapitre et présente des recommandations en matière de recherches et d'applications futures.

Importance pour la défense et la sécurité

L'ARC s'est engagée à accroître l'efficacité opérationnelle de ses équipages, ce qu'elle peut notamment réaliser grâce à de nouvelles technologies conçues pour améliorer la sélection, la formation ou le rendement opérationnel de son personnel. La présente analyse montre que la formation et le rendement opérationnel des pilotes militaires peuvent être améliorés au moyen d'applications de suivi des mouvements oculaires. Le présent rapport a été rédigé à la demande du DSPA et constitue le produit livrable final du projet 03rm de RDDC.

Table of contents

Abstract i
Significance to defence and security i
Résumé ii
Importance pour la défense et la sécurité ii
Table of contents iii
List of figures iv
List of tables v
1 Introduction 1
2 Eye movements and visual attention 2
3 Eye movements in the cockpit 6
3.1 Eye movement measures 7
3.1.1 Proportional dwell time 7
3.1.2 Durations of fixations and dwells 8
3.1.3 Frequency of fixations and dwells 9
3.1.4 Saccade Length 10
3.1.5 Dwell sequences and patterns 11
3.1.5.1 Transition matrices 11
3.1.5.2 Global pattern metrics 12
3.1.6 Microsaccades 13
3.1.7 Pupil diameter 13
3.1.8 Blinks 14
3.1.9 Perceptual span 14
3.1.10 Event-related analysis 15
3.2 Models 16
3.3 Other applications 17
3.3.1 Usability evaluation 17
3.3.2 Field-of-view requirements 17
3.3.3 Qualitative feedback during training 18
3.3.4 Aircraft inputs 18
4 Summary 19
5 Future research 20
References 23
Annex A Summary table 33
Annex B Eye-tracking technology primer 43
List of symbols/abbreviations/acronyms/initialisms 46

List of figures

Figure 1 A schematic showing a series of fixations and dwells separated by saccades. 3

Figure 2 An illustration of the information in the image captured by retinal photoreceptors, for two successive fixations (left panel: fixation on airspeed indicator, right panel: fixation on horizontal situation indicator). Due to the distribution of photoreceptors over the retina, high resolution colour visual information is extracted from the central few degrees about the point of gaze. Resolution and colour sensitivity drop off steeply with increasing eccentricity from the point of gaze. Note that this figure is a qualitative illustration only and does not depict actual filtering values (both in colour saturation and blur) at each retinal eccentricity. Furthermore, the visual experience of the viewer is the result of extensive post-retinal processing. 4

List of tables

Table A.1 Chronological summary of the empirical studies that were considered in this review. 33

This page intentionally left blank.

1 Introduction

The modern aircraft cockpit presents a complex visual scene. The aviator's field of view is occupied by a dense mosaic of instruments and displays that convey information about the status of the aircraft. By observing these information displays the aviator can assess the state of the aircraft, including its movement vector, flight and control systems, automation systems (e.g., autopilot), navigation and communications systems. In military aircraft there may be additional displays containing sensor (e.g., radar, infrared camera) and weapon information. As the flight progresses, displayed information changes in real-time and must be closely monitored. This presents a challenge to human operators, who process information from the visual field in a serial manner with eye movements determining each subsequent location for processing. Accordingly, eye movement recordings may provide a window onto the aviator's processing of information from cockpit displays.

Research into the relationship between eye movements in the cockpit and pilot performance began in the years following the Second World War. Presently we review this domain of research to date. The paper begins with a brief overview of relevant characteristics of human eye movements and visual attention followed by a review of empirical research examining eye movements during aircrew performance. The review is organized by eye movement measures and focuses on the relationships of each measure to pilot expertise, mental workload, and situational awareness. Details about the methodology for each citation are collected for quick reference in Annex A and a primer on current eye-tracking technology is contained in Annex B. Finally, suggestions are provided for future research in this domain.

2 Eye movements and visual attention

Humans make various eye movements including saccades, smooth pursuit movements, vergence eye movements, and vestibulo-ocular reflex movements (for a review see Rayner, 1998, 2009). *Saccades* are rapid, ballistic eye movements that direct the point of gaze to another area of the visual field. Processing of information from the visual field is suspended during saccades (known as saccadic suppression). S*mooth pursuit* eye movements are made in order to continuously align the point of gaze with a target moving across the visual field (e.g., a passing aircraft). In contrast to saccades, during smooth pursuit movements the visual perception of the target stimulus continues in order to update eye velocity and maintain pursuit (Masson and Stone, 2002). *Vergence* eye movements serve to align the eyes when focusing on different depth planes. For example, when the pilot focuses on the near field (e.g., instrument panel) his/her eyes will converge, and when focusing on a distant object through the windscreen they will diverge. Finally, *vestibulo-ocular* movements occur automatically to compensate for movements of the head. These reflex eye movements allow the visual system to maintain a constant point of gaze within the visual field amid frequent head movements.

Saccades are particularly important in the context of information processing of cockpit instruments and displays. While vision is suppressed during a saccade, at the end of a saccade the eyes fixate a point in the visual field and a relatively stable image is projected onto the retina. It is during these *fixation* events that visual information is extracted from the fixated area. During tasks such as reading, scene perception, and visual search, fixation durations range from ~ 40 ms to over 800 ms, with mean fixation durations typically falling between 200 ms and 400 ms depending on the viewing task (Rayner, 1998). The link between fixations and cognitive processing has been the subject of decades of research, with consensus emerging that fixation durations depend partly on cognitive processing of the fixated material. For example, during reading, fixations that land on low-frequency (i.e., relatively rare) words have longer duration than fixations on high frequency words, and this is thought to be due to differences in the time required to identify the words (for a review see Reingold, Reichle, Glaholt & Sheridan, 2012). By analogy to the context of cockpit instrumentation, instruments or displays that contain information that is relatively difficult to read might be expected to produce longer fixation durations.

A viewer might also make multiple fixations to a particular region of the visual field. For example during scene perception, a viewer might inspect a particular area of the scene (e.g., a single cockpit instrument) in a series of small saccades before moving to another area of the scene in a larger saccade. A group of fixations on a particular piece of information within a display are sometimes referred to as a *dwell* (e.g., Glaholt & Reingold, 2009; for an illustration see Figure 1).

Figure 1: *A schematic showing a series of fixations and dwells separated by saccades.*

Furthermore, the area from which high detail (i.e., high spatial frequency) information is captured during a fixation is relatively narrow. This is a consequence of the distribution of photoreceptors over the retina: the density of cone receptors that sense detailed (i.e., high spatial frequency) colour information is high in the central few degrees of the retinal visual field (known as the *fovea*) and drop off steeply toward the periphery (Yamada, 1969). Rod photoreceptor density increases towards the periphery, but rods provide low-resolution monochrome visual information and are optimized for vision at night as well as the detection of motion (Takeuchi & Valois, 2009). Thus humans are only sensitive to high resolution tri-chromatic visual information within the central few degrees about their point of gaze (see Figure 2). Functionally this means that to process detailed visual information such as the information displayed by aircraft instruments, it must be fixated precisely so that its image is cast upon the fovea of the retina. The primary function of saccades is to direct the fovea to another area of the visual field.

Figure 2: *An illustration of the information in the image captured by retinal photoreceptors, for two successive fixations (left panel: fixation on airspeed indicator, right panel: fixation on horizontal situation indicator). Due to the distribution of photoreceptors over the retina, high resolution colour visual information is extracted from the central few degrees about the point of gaze. Resolution and colour sensitivity drop off steeply with increasing eccentricity from the point of gaze. Note that this figure is a qualitative illustration only and does not depict actual filtering values (both in colour saturation and blur) at each retinal eccentricity. Furthermore, the visual experience of the viewer is the result of extensive post-retinal processing.*

For a given viewing task, the area of the visual field over which a viewer extracts useful information within an eye fixation is known as the *perceptual span* (Rayner, 1975; Rayner, Slattery, & Bélanger, 2010). Researchers have developed techniques to estimate the size of the perceptual span for specific viewing tasks and for individual viewers. For example, the *gaze-contingent window* manipulation involves monitoring eye movements and updating the viewer's display contingent on gaze position such that a window of a certain size is continually centered at the viewer's point of gaze (McConkie & Rayner, 1975). Visual information in the display is visible within the window and is masked outside of it. The logic of this manipulation is that if masking part of the visual field causes task performance to be impaired, then it can be inferred that the viewer was processing and using information from that part of the visual field in support of performance. By manipulating the size of the gaze-contingent window, researchers can determine the point at which the presence of the window impairs performance, and thus infer the size of the perceptual span of an individual for a given task. Prior research has shown that the size of the perceptual span increases with task-specific expertise (Reingold & Charness, 2005).

In addition to cognitive processing of the fixated visual information, during an eye fixation the human visual system must also select a target for the upcoming saccade. The selection of saccadic targets from the visual field is also a subject of much research and theory (Findlay & Walker, 1999; Zelinksy, 2008). In particular, it is known that saccadic targeting is influenced by low-level salience of image features in the visual field (so-called *bottom-up* effects on saccadic selection). For example, areas of the visual field with high contrast or high edge density tend to attract

fixations. Saccadic selection is also influenced by higher level task goals, strategies, or prior knowledge (*top-down* factors; e.g., Yarbus, 1967). In either case, the selection of a target for the upcoming saccade is likely to depend on processing of visual information outside of the fovea (in the parafoveal or peripheral visual field).

Human viewers are also capable of selectively devoting processing resources to areas of the visual field without moving their eyes. This faculty, known as covert visual attention, was first demonstrated by Posner and colleagues (Posner, 1980; Posner, Snyder, & Davidson, 1980) using a spatial cueing paradigm. In this paradigm, the viewer keeps their eyes fixated on a central point, and is required to detect targets that appear left or right of the central point. When the frequency of the target presentation is biased towards one side of the display, subjects are faster to detect that target. This demonstrates that sensitivity to visual stimuli in different areas of the visual field can be biased through the top-down influence of visual attention. Since the Posner et al. findings, there has been a great deal of research on the relationship between the point of gaze and the focus of visual attention, however the consensus is that during normal behaviour, eye movements and visual attention are tightly coupled (see Findlay, 2009). More specifically, it is thought that a movement of gaze to a specific point in the visual field is preceded by a movement of visual attention to that location. Therefore under natural viewing the spatial distribution of eye movements corresponds closely to the distribution of visual attention and also the distribution of visual processing over the visual field.

Several other oculomotor activities are potentially relevant to aircrew flight performance: blinks, microsaccades, and changes in pupil size. The *blink* is a ubiquitous oculomotor action that lubricates and clears the corneal surface. However in laboratory tasks, blinks have been shown to correlate with mental workload (Ahlstrom & Friedman-Berg, 2006; van Orden, Limbert, Makeig, & Jung, 2001) and fatigue (Morris & Miller, 1996). In addition, percent eye closure has been used as a measure of fatigue and alertness (Dinges & Grace, 1998; Wierwille, 1999). *Microsaccades*, part of a class of so-called fixational eye movements, occur while the eye is in fixation. One function of microsaccades is to prevent adaptation to the visual stimulus: if a fixed image is cast across the retina (known as retinal stabilization), the percept of the projected image fades until the image is shifted again. Microsaccades shift the retinal image slightly and prevent this adaptation. Recent research has also suggested a link between microsaccades and mental fatigue (Di Stasi, McCamy, Catena, Macknik, Cañas, & Martinez-Conde, 2013) in laboratory tasks. Finally, the *pupil* is the aperture in the cornea through which light passes before hitting the retina. The diameter of the pupil is adjusted automatically through smooth muscle movements in response to ambient luminance (for a review see Watson & Yellott, 2012). However, there is also evidence that pupil diameter increases with increasing cognitive workload (Ahlstrom & Friedman-Berg, 2006; de Greef, Lafeber, van Oostendorp, & Lindernberg, 2009), and that pupil measures correlate with sleep deprivation (Morad, Lemberg, Yofe, & Dagan, 2000; Russo, Thomas, Thorne, Sing, Redmond, Rowland, Johnson, Hall, Krichmar, & Balkin, 2003).

3 Eye movements in the cockpit

A central appeal to eye movement monitoring in the cockpit is the possibility of a window onto the aircraft operator's cognitive state. The cognitive state of the operator is obviously important for safe and effective aircraft operation. For example, a pilot that is overloaded with concurrent cockpit tasks might be more likely to miss a critical warning indicator in the cockpit display, or perhaps fail to notice the appearance of another aircraft in nearby airspace. Mental workload has been defined previously as the ratio between the operator's task demands and the operator's cognitive capacity (Kantowitz, 1987). Performance variables that are derived from aircraft control inputs (e.g., throttle, flight stick) can provide some indication of operator overload, such as erratic control behaviour or deviation from course. But even when an operator's task demands exceed their capacity and he or she is overloaded, there might not be overt behavioural signs. For example, lapses in situational awareness (see Jones & Endsley, 1996; van Dijk, van der Merwe, & Zon, 2011) might not result in changes to aircraft control inputs. Consider a situation in which a novice is overloaded by a particular flight task requirement while an expert, due to efficiency in processing, is not overloaded and thus maintains high situational awareness. Accordingly, the expert might notice visual cues that are missed by the novice. However, if these cues do not require action, the expert's 'noticing' of those cues would leave no trace in performance measures and therefore it would be difficult to distinguish between the two pilots. Experts are also known to be more selective in the information they process, focusing their attention on information relevant to the decision or problem at hand (see Schriver, Morrow, Wickens, & Talleur, 2008). However, a pilot's selectivity in processing cockpit information may only be weakly reflected in their operation of the aircraft.

Eye movements have the potential to provide additional, and possibly more direct, measures of pilots' information processing in the cockpit. Based on the what is known about the functions of eye movements in other domains (e.g., reading, scene perception, visual search), eye movements should provide indications of a) the information that is sampled by the operator over a given period (e.g., the distribution of fixation locations) and b) the time that it takes to process (e.g., fixation durations). Both of these aspects of eye movements are potentially sensitive to operator workload and expertise. In particular, the expert's pattern of visual attention is likely to reflect their efficiency in selecting and processing context-relevant cockpit information. In addition, eye movement recordings might capture 'latent' differences between experts and novices under high and low mental workload. Together with performance measures, eye movements have the potential to yield a more complete picture of the operator's workload, expertise and situational awareness.

Not surprisingly, aircrew eye movements have been a topic of great interest in civilian and military aviation research since the Second World War. In an early application, Fitts and colleagues (Fitts, Jones, & Milton, 1949; 1950) sought to determine the optimal arrangement of instruments within the cockpit instrument panel. By video recording and manually classifying eye movements of pilots while they flew a C-45 aircraft, Fitts and colleagues were able to provide design recommendations that resulted in the familiar "T"-shape instrument arrangement and cross-check (see also Senders, 1964; 1966). Importantly, this early research identified eye movement measures that have been pursued in subsequent research: fixation duration, fixation frequency, and fixation pattern. In the years since this seminal work by Fitts and colleagues, a group of scientists at National Aeronautics and Space Administration (NASA) Langley Research

Center conducted a series of studies aimed at identifying expertise and workload effects from eye movements in the cockpit (e.g., Jones, 1985; Senders, 1966; Tole, Stephens, Vivaudou, Harris, & Ephrath, 1983; for a review see Harris, Glover, & Spady, 1986). Since the early 1990s, improvements to video-based eye tracking technology have provided cost-effective, automatic recording of gaze position with high spatio-temporal precision (see Annex B for a primer on eye tracking technology), allowing for measurement of saccades, pupil size, blinks, and even fixational eye movements (e.g., microsaccades). This has resulted in a third wave of research emerging from a multiple labs.

The majority of studies in this domain have used video-based eye-trackers to monitor pilots' eye movements while they complete flight tasks either in a simulator or in a real aircraft. Flight tasks might include take-off, landing, level flight, turn and maneuver, conduct cockpit procedures, or detect events or anomalies that appear within the instruments or out the window. Studies that examine mental workload in the cockpit typically compare flight tasks that differ in difficulty (e.g., compare landing vs. level flight), or else difficulty is manipulated within a task (e.g., add turbulence), or through the introduction of a secondary task that varies in difficulty (e.g., an auditory counting task). Studies that examine pilot expertise typically compare groups of subjects with different experience levels (e.g., students vs. instructors; differences in total hours in flight). Studies on situational awareness typically test for subjects' awareness of in-flight events such as a system failure, a change in aircraft automation settings, or the appearance of objects in nearby airspace.

Because the number of empirical papers on eye movements in the cockpit is large, it would be onerous to review them in a detailed chronological fashion. Instead we have organized the review by eye movement variable, summarizing for each variable the connection to pilot performance, workload, expertise, and situational awareness. In order to help illustrate the body of work in this domain we have created a table (see Annex A) that captures for each empirical study, the eye tracking system used, the nature of the flight task (e.g., simulator, take-off and landing task), the research factors of interest (e.g., workload, expertise), and the eye movement measures obtained. The following sections review eye movement measures in detail.

3.1 Eye movement measures

3.1.1 Proportional dwell time

A common approach to analyzing pilot gaze behaviour is to divide the cockpit into relevant areas and compute total viewing time (i.e., dwell time) for each area. A separate area is typically defined for each cockpit instrument, display, and for the windscreen (out-the-window; OTW). The total dwell time for each area is then divided by the sum dwell time across all areas in order to derive the proportional dwell time (PDT) for each instrument (often reported as a percentage). One clear finding from prior research is that PDT is not equal across areas of the cockpit and some areas receive more dwell time than others (e.g., Alexander & Wickens, 2006; Anders, 2001; Hayashi, 2004; Hayashi, Ravinder, McCann, Beutter, & Spirkovska, 2009; Krebs, Wingert, & Cunningham, 1977; Matessa & Remington, 2005; Mumaw, Sarter, & Wickens, 2001; Spady, 1978; Wierwille, Rahimi, & Casali, 1985). Areas for which the displayed information changes frequently (e.g., the primary flight display; Mumaw, et al., 2001) or that is of critical importance to the flight task (e.g., van Dijk, et al., 2011), are likely to receive a larger proportion of the total

dwell time. Svensson et al. (1997) found that the amount of dwell time spent heads-up (i.e., OTW) depended on the amount of information presented heads-down (in this case, a tactical display); more information heads-down was associated with more dwell time heads-down. In addition, the proportional dwell time for instruments can vary across phases of flight (Brown, Vitense, Wetzel, & Anderson, 2002; Diez, Boehm-Davis, Holt, Pinney, Hansberger, & Schoppek, 2001; Colvin, Dodhia, & Dismukes, 2005; Huemer, Hayashi, Renema, Elkins, McCandless, & McCann, 2005; Mumaw, et al., 2001; Spady, 1978), presumably due to the changes in requirements for information from various instruments. Perhaps surprisingly, few studies have reported evidence for an effect of mental workload on proportional total dwell time across instruments. However, Spady (1978) found that the pattern of PDT across instruments changes with air turbulence (see also Krebs, et al., 1977).

Recent studies by Li and colleagues (Li, Chiu, & Wu, 2012; Li, Chiu, Kuo, & Wu, 2013) showed that the pattern of PDT over instrument panels depended on pilot expertise (junior vs. senior pilots). Also, Schriver et al. (2008) found that expert pilots have higher PDT on relevant cockpit information during a failure event compared to pilots with less experience. Note that the PDT analysis can also be applied to areas defined within the OTW region of the cockpit. For example, Kim, Palmisano, Ash, and Allison (2010) compared PDT across different parts of the windscreen during visual-flight-rule (VFR) landing and were able to distinguish between novice and experienced pilots' patterns of OTW dwell time. Colvin, Dodhia, Belcher, Dismukes, Poly, & Obispo (2003) also reported individual differences in the PDT directed OTW, with individuals varying between 32% and 61% percent dwell time out-the-window. Other researchers have measured proportional dwell time OTW in order to assess the impact of novel cockpit instruments on situational awareness in nearby airspace (Cote, Krueger, & Simmons, 1985; Johnson, Wiegmann, & Wickens, 2006; Oseguera-Lohr & Nadler, 2004; Wickens, Alexander, Thomas, Horrey, Nunes, Hardy, & Zheng, 2004; Willams, 2002).

Total dwell time, from which PDT is derived, is the cumulative duration of all dwells on an area. As such it will be determined by the duration, as well as the frequency, of individual dwells. Furthermore dwells are composed of individual fixations, and hence dwell duration depends on the number of fixations and also the duration of individual fixations. These various measures are discussed in the following sections.

3.1.2 Durations of fixations and dwells

As was discussed in section 2.1, fixation durations are partly determined by the time required to cognitively process the fixated material. Consistent with this, Kramer, Tham, Konrad, Wickens, Lintern, Marsh, Fox, and Merwin (1994) showed that expert pilots have shorter fixation durations on cockpit instruments than novices. In addition, several studies have documented differences in fixation duration as a function of the area (e.g., instrument/display) fixated within the cockpit (Gainer & Obermayer, 1964; Senders, 1966; Spady and Harris, 1981), suggesting that cockpit display areas might differ in either a) the ease with which information is encoded from them (see Gainer & Obermayer, 1964) or b) the bandwidth of information contained in them (see Senders, 1966). During system failures, fixation durations have been shown to increase for instruments that display critical information (Itoh, Hayashi, Tsukui, & Saito, 1990). Diez, Boehm-Davis, Holt, Pinney, Hansberger, and Schoppek (2001) also showed a relationship between fixation duration and situational awareness, where long duration fixations on relevant cockpit instruments correlated with high situational awareness. The effect of flight task and mental workload on

fixation duration is less empirically clear. Ellis (2009) failed to find a significant relationship between fixation duration and workload. Jessee (2010) found no effect of flight task on fixation duration, but found that the variability in fixation duration changed across tasks.

A dwell is defined as a set of one or more consecutive fixations on the same area, instrument or display[1]. Dwells are multi-fixation processing epochs, and their durations have been shown to be sensitive to pilot expertise and mental workload. In particular, expert pilots tend to have shorter dwell durations than novices (Fitts et al., 1949; Bellenkes, Wickens, & Kramer, 1997; Kasarskis Stehwien, Hickox, Aretz, & Wickens, 2001; Sullivan, Yang, Day, & Kennedy, 2011; for a null finding see Yang, Kennedy, Sullivan, & Fricker, 2013), reflecting the efficiency with which experts process cockpit display information. Dwell durations also differ depending on the instrument they are directed to (Fitts, et al., 1949; Sarter, Mumaw, & Wickens, 2011; Spady, 1978; Tvaryana, 2004), and over different phases of flight (Katoh, 1997). In contrast to the weak evidence in the context of fixation duration, for dwell duration there are several demonstrations of a correlation with task difficulty and workload. Harris, Glover, and Spady (1986) showed that increased mental workload tended to increase the duration of dwells. This increase was due to an increase in the number of fixations that composed the dwell rather than a lengthening of the component fixations. Spady (1978) and Dick (1980) presented evidence that dwell durations are longer for manual flight control than with automation. Causse, Baracat, Pastor, & Dehais (2011) found that dwell durations increased during instrument landing when there was high uncertainty about the safety of landing. Finally, Tole et al. (1983) reported that dwell durations on instruments increased with mental workload imposed by an auditory secondary task, but that this increase was greater for novices than experts, indicating an interaction between expertise and processing time (see also Robinski & Stein, 2013).

3.1.3 Frequency of fixations and dwells

In addition to the duration of individual fixations and dwells, one can conduct the complementary analysis the frequency that a particular cockpit area is fixated (or dwelled upon) within a flight segment or epoch. Consistent with the measures discussed previously, it is clear that cockpit areas are not fixated with equal frequency (Chuang, Nieuwenhuizen, & Bülthoff, 2013; Fitts et al., 1949, 1950; Hayashi et al., 2009; Tvaryana, 2004; van de Merwe, Dijk, & Zon, 2012). Van de Merwe et al. (2012) found during a malfunction problem solving task, fixation rates on various instruments were related to their problem-relevance. Colvin, Dodhia, & Dismukes (2005) found that even under visual flight rules (VFR) when out-the-window viewing is considered critical, the instrument panel was fixated more frequently than the OTW area (for a similar finding see Wickens et al., 2004).

[1] The terms 'dwell' and 'fixation' are sometimes used interchangeably but for the present purposes they must be distinguished. The term 'fixation' describes a physiological event: a pause in gaze position bounded by movements of the eye (e.g., saccades or blinks). 'Dwell' refers to a visit to a particular location in the display, such as an instrument. In viewing an instrument, the pilot may make one or more consecutive fixations on the instrument before leaving it to gaze elsewhere. The set of consecutive fixations on the instrument constitute a dwell, and hence dwells, unlike fixations, are determined by the fixated content. The confusion of these terms may be due, in part, to the historical technological challenge in identifying physiological fixations, which are typically bounded by high-velocity eye movements: eye trackers that have a slow sampling rate or a low spatial precision might not be able to reliably detect rapid, small saccades, and thus the analysis must be based on dwells bounded by areas of the visual field.

Fixation and dwell frequency are also sensitive to pilot expertise. Bellenkes, Wickens, & Kramer (1997) found that some cockpit instruments were dwelled upon more frequently by experts than novices, and Ottati, Hickox, & Richter (1999) found that experienced pilots made more fixations overall than novices. Similarly, Kasarskis et al. (2001) found that expert pilots made more frequent fixations than novices, and that high fixation frequency was associated with ‘good’ landing performance in both groups.

The link between fixation frequency and pilot mental workload has not been well established, though there is some evidence of an effect of task demands on fixation frequency. In particular, Spady (1978) examined eye movements during simulated landing approach under instrument flight rules (IFR). Simulated turbulence was found to increase dwell frequency overall, and that manual throttle control produced a reduction in dwell frequency and increase in dwell duration on the attitude indicator compared to autopilot. Brown, Vitense, Wetzel, and Anderson (2002) also found that when flying manually, pilots fixated the primary flight display less frequently but for longer durations compared to when flying on autopilot.

It is important to note that some reports of fixation or dwell frequency effects show complementary effects on fixation and dwell duration. For example, Kasarskis et al. (2001) reported that experts made more fixations than novices, and also that experts made shorter fixations than novices. This suggests a potential confound between fixation duration and frequency: if a subject makes relatively short fixations, they will also produce more frequent fixations over a given interval. Hence a caveat for interpreting overall (i.e., global) fixation and dwell frequency effects is that they may be a consequence of short overall fixation and dwell durations.

3.1.4 Saccade Length

Saccade length (also referred to as saccade amplitude) is the distance between fixated points in the visual field and corresponds to the angle through which the eye travels during a saccade. This variable is not often considered in the context of cockpit gaze behaviour, yet prior research has produced some evidence that saccade length is sensitive to pilots’ cognitive state. For example, Krebs et al. (1977) found that increased mental workload (via the introduction of air turbulence) caused pilots to produce shorter saccades. The authors suggested that this might reflect a tendency for pilots to focus on a few instruments during high workload conditions. Katoh (1997) also analyzed saccade amplitudes during simulated flight, and found that saccade amplitudes depended on flight task requirements: instrument-based tasks promoted short saccades (likely between instruments) where ‘scenery’-based (i.e., OTW) produced larger saccades (likely between OTW regions). Jesse (2010) also found differences in saccade length as a function of flight task. Schnell, Kwon, Merchants, and Etherington (2004) evaluated a novel cockpit display (a so-called “synthetic vision system”) and found that it caused a reduction in total saccade length, which may have reflected pilots’ focus on the new display item. Finally, Dahlstrom and Nahlinder (2009) computed a total eye movement energy which sums the vertical and horizontal eye movements over a flight period. The authors found that eye movement energy differed between simulated and actual flight, but interpreted this as being due to the presence of the OTW view in real flight as opposed to the instruments-only display in the simulator.

3.1.5 Dwell sequences and patterns

A central goal of early research on eye movements in the cockpit was to identify higher-order patterns of eye movements. More specifically, there have been two main approaches to characterizing dwell pattern in the cockpit. The first approach analyzes dwell transitions between cockpit instruments, displays, and area (e.g., Fitts et al., 1949, 1950; Senders, 1966; Jones, 1985). The second approach attempts to derive global pattern metrics such as *entropy* (e.g., Tole et al., 1983) and *nearest-neighbour index* (Di Nocera, Camillia, & Terenzi, 2007). These two approaches are discussed in turn.

3.1.5.1 Transition matrices

Fitts and colleagues (Fitts et al., 1949, 1950) sought to determine the optimal arrangement of cockpit instruments from the perspective of pilot instrument scan. To do this, they measured the 'link' between each instrument by computing the frequency of dwell transitions between two instruments (i.e., both directions) and dividing this value by the total number of transitions. Fitts et al. found that some pairs of instruments had higher link values than others and suggested that this information could be used to evaluate instrument arrangements; optimal arrangements would put 'linked' instruments close together. This work influenced the standard instrument panel arrangement used in most Western aircraft (Jones, 1985).

Subsequent research considered a more general treatment of dwell pattern over cockpit areas by deriving transition matrices (Carbonell, 1966; Senders, 1966; for a review see Harris et al., 1986). In particular, the value of a given cell in the transition matrix is the probability of a dwell being directed to a particular cockpit display or area, contingent upon the location of the previous dwell. Extending the transition matrix to three dimensions, the value of a given cell in the matrix is the frequency of dwells on an area contingent on the location of the previous two dwells. These contingency analyses have been referred to as Markov chains (Norris, 1998). The previous example would constitute a 2nd order Markov chain. A zero-order Markov chain would be the simple probability of a dwell being directed a cockpit display (i.e., with no prior history).

There has been some controversy among researchers about whether or not pilots actually have consistent higher-order dwell patterns per se, or if their sequential dwell patterns could be adequately described by simple dwell frequency data (e.g., Senders, 1966). If the latter is true, it would mean that pilots randomly choose subsequent instruments to dwell upon while drawing from the simple probability distribution of dwells for each instrument. Further work has contested this (Tole et al., 1983; Ellis & Stark, 1981; Ellis, 1982), arguing that deterministic scan patterns are difficult to detect and require sensitive statistical techniques (see Seeberger & Wierwille, 1976). Indeed, higher-order dwell sequences will be intrinsically more difficult to observe than lower-order sequences from a statistical perspective (e.g., every sequence of three dwells contains one 2rd-order chain, two 1nd-order chain and three zero-order probabilities), making it more likely to achieve statistical power for lower-order sequences than higher-order ones.

Nevertheless, studies on pilots in the cockpit have identified systematic dwell patterns from transition matrices. For example, dwell patterns captured in transition matrices have been shown to depend on the flight maneuver being carried out (Gainer and Obermayer, 1964; Jones, 1985). Dependence of dwell pattern on flight context (i.e., task, situation) might actually introduce difficulty for analyses designed to detect overall dwell patterns; if a pilot rapidly switches

between two higher order dwell transition patterns, the patterns may be blurred or lost when analyzed together. To address this possibility, Hayashi (2004) proposed a Hidden Markov Model approach which includes 'hidden' Markov states that correspond to different flight tasks, and allows for partially overlapping dwell patterns between tasks. This has subsequently been applied modeling the dwell patterns of space shuttle crew (Hayashi, Beutter, & McCann, 2005).

Dwell patterns have also been linked to expertise and workload. For example, Harris et al. (1986) observed changes in dwell pattern under atmospheric turbulence, and also when automation is engaged. In a study by Tole et al. (1983), pilots of varying skill completed a simulated landing under instrument flight rules while conducting an auditory secondary task that varied in difficulty. The authors computed transition matrices and compared the frequency of the top ten most frequent 3rd-order transition sequences for each pilot. Interestingly, the frequency of the top ten sequences decreased for each pilot as mental workload imposed by the secondary task increased, suggesting that the dwell pattern became more variable as workload increased. Importantly, for skilled pilots this effect was reduced, indicating that their dwell pattern was more robust to mental overload than for less experienced pilots.

Individual differences in dwell pattern have been observed. In particular, Dick (1980) found dwell patterns that were specific to individual pilots (see also Chuang et al., 2013). Also, Kramer et al. (1994) found that higher-order Markov chains were useful for distinguishing between student and expert pilots, but this was because there was more variability in dwell pattern among the experts than the students. Indeed, one of the possible difficulties with measures of dwell pattern is the potential for robust individual differences in the manifestation of expertise in dwell pattern; experts might have idiosyncratic, yet equally effective, cockpit eye movement patterns.

3.1.5.2 Global pattern metrics

Given that dwell patterns reflect order in a pilot's eye movement behaviour during flight, one approach to detecting the presence of dwell patterns is to measure the extent to which the dwell sequence is disordered or random. Entropy is an information theory measure (Shannon & Weaver, 1949; for a review see Verdü, 2000) that captures the degree of disorder or randomness in a sequence. By analogy to the cockpit, dwell sequences over cockpit areas that are less ordered and more random will have high entropy. Ephrath, Tole, Stephens, and Young (1980) examined eye movements of test pilots during simulated landings and manipulated mental workload both through task difficulty and via a secondary task, and reported a monotonic increase in entropy with increasing mental workload. Two more recent studies (van Dijk et al., 2011; van de Merwe et al., 2012) found that entropy increased following a cockpit instrument failure, conditions that presumably would entail increased mental workload.

A measure known as approximate entropy (ApEn) was developed by Pincus (1991). ApEn is an entropy measure that is better suited to situations in which there is a lot of measurement noise in addition to underlying signal noise (e.g., the recording of biological signals). McKinley, McIntire, Schmidt, Repperger, and Caldwell (2011) investigated changes in ApEn associated with fatigue in pilots. Eye movements were recorded while pilots conducted various flight tasks. McKinley et al. found a decrease in ApEn at high fatigue levels when pilots performed an out-the-window target identification task. However in a simulated Unmanned Aerial Vehicle landing task, there was no significant difference in ApEn as a function of fatigue.

The entropy measure considers the sequence of dwell locations only, and discards dwell duration. However, as was discussed earlier, mental workload and expertise have both been shown to affect dwell duration. In order to incorporate dwell duration, a measure of entropy rate has been suggested: the entropy of a dwell sequence divided by the duration of that sequence (Tole et al., 1983). Tole et al. (1983) found that increased mental workload imposed by a secondary auditory task tended to lower entropy rate, but this pattern was only present for novice pilots and not for experts. Harris et al. (1986) also report that eye movement entropy rate decreases as pilot mental workload increases, and also that novice pilots exhibit a low entropy rate early in training, but as training progresses their entropy rate increases to match that of experts. Itoh et al. (1990) examined entropy rates across different instrument panels and found that the entropy rate for a panel with an integrated information display produced a smaller entropy rate, likely due to the clustering of gaze around the integrated display panel.

With regards to mental workload there seems to be some discrepancy between the findings using the entropy rate measure and those reporting entropy: while entropy was found to increase with mental workload, entropy rate appears to decrease with mental workload. This suggests that dwell duration is an important consideration when computing the degree of randomness in instrument viewing, however these differences might also be due to variations in the way that mental workload is manipulated between experiments. Further research in this area might seek to resolve this discrepancy.

Another global pattern measure known as Nearest Neighbour Index (NNI) was introduced by Di Nocera et al. (2007). Nearest neighbour index is a measure of spatial clustering, and is computed by summing the distances of each fixation to its nearest neighbour and dividing this sum by the average distance between fixations derived from a uniform random distribution. Values less than one indicate departures from a random spatial distribution. Note that while NNI detects departure from spatial randomness, it not sensitive to the degree of randomness in the sequence of fixation or dwell locations (c.f. entropy). Di Nocera et al. (2007) found that NNI varied across phases of simulated IFR flight, showing the least random (most clustered) distribution of fixations during cruising flight and the most random (least clustered) distribution during take-off and landing which are expected to have the highest mental workload.

3.1.6 Microsaccades

At the time of writing, no study to date has examined aircrew microsaccades. However there is promising research from non-flight domains that link microsaccade rate to mental fatigue (Di Stasi, McCamy, Catena, Macknik, Cañas, & Martinez-Conde, 2013). In particular, Di Stasi et al. monitored subjects' eye movements during a simulated air-traffic control task and observed a robust reduction in microsaccadic peak velocity as time-on-task increased (over four 30 minute sessions). This is an encouraging finding and the potential application to aircrew eye movements during flight should be explored.

3.1.7 Pupil diameter

Several studies have investigated changes in pupil diameter during aircrew performance (Cheung & Hofer, 2003; Krebs et al., 1977; Previc, Lopez, Ercoline, Daluz, Workman, Evans, & Dillon, 2009; Wierwille et al., 1985; Wilson, Caldwell, & Russell, 2007). In an early demonstration, Krebs et al. (1977) found that pupil diameter increased for flight segments that were subjectively

rated as high workload. Wilson et al. (2007) found a similar effect in the context of an unmanned aerial vehicle (UAV) sensor feed monitoring task. However, Previc et al. (2009) found no significant differences in pupil diameter across flight segments, and Wierwille et al. (1985) found pupil diameter to be relatively insensitive to pilot mental workload. Recently, Li et al. (2013) reported that senior pilots (compared to juniors) had increased pupil diameter when viewing relevant cockpit displays, suggesting a possible relationship between pupil diameter and expertise. As an aside, Cheung and Hofer (2003) induced a pitch illusion in pilots during simulated flight and found that pupil diameter increased during the spatial disorientation that followed. Hence overall there is a mixed pattern of findings with pupil diameter. One of the caveats to using pupil diameter as a measure of pilots' cognitive state is that it correlates strongly with other eye movements variables (e.g., dwell duration; see Krebs et al., 1977). Pupil diameter also varies autonomously with ambient illumination (Watson & Yellot, 2012), which is difficult to control in the cockpit but can be controlled to some extent within flight simulators.

3.1.8 Blinks

Blinks have been considered as possible indicators of pilot cognitive state (Hankins & Wilson, 1998; Krebs et al., 1977; McKinley et al, 2011; Veltman & Gaillard, 1996; Wierwille et al., 1985; Wilson, Fullenkamp, & Davis, 1994; Wilson, 2002; Wilson & Fisher, 1991). Wilson (2002) suggested that blink rates might offer a measure of 'visual demand', finding that blink rate decreased for visually demanding flight segments (IFR vs. VFR, landing vs. cruise). This general pattern has been replicated across several studies comparing flight segments of varying difficulty (Hankin & Wilson, 1998; Veltman & Gaillard, 1996; Wilson, Fullenkamp, & Davis, 1994) as well as mental workload imposed by a secondary task (Wierwille et al., 1985). There is also some evidence for a relationship between blinks and time-on-task. Stern and Bynum (1970) measured blink rate in helicopter pilots during actual flight, and found that blink rate decreased with time on task, and more so for novices than for experts. McKinley et al. (2011) manipulated time-on-task to investigate fatigue effects during aircrew tasks (OTW target identification and UAV control), and computed *percent eye closure,* the proportion of time that the eye was closed during a task. While the data from McKinley et al. (2011) show some indication that percent eye closure increases with time-on-task, the measure was also noisy and the authors concluded that it may not be a sufficiently robust signal with which to track fatigue. Blink rate and other associated variables (such as blink duration, inter-blink interval) have been used to classify flight segments (Wilson & Fisher, 1991).

3.1.9 Perceptual span

Only one study has used eye tracking to study the perceptual span of pilots in the cockpit. Fox, Merwin, Marsh, McConkie, and Kramer (1996) tested the hypothesis that compared to novices; expert pilots would be able to extract visual information from a larger area about the fixated point. This could allow experts to process information from instruments that are not directly fixated (i.e., in the parafoveal or peripheral visual field). In order to test this hypothesis they employed a gaze-contingent display manipulation whereby instrument information was only visible when gaze was directed to the instrument, and masked when gaze was directed elsewhere. Therefore in the gaze-contingent viewing mode the parafoveal and peripheral 'preview' of the instrument status is denied. The logic of this manipulation is that if a pilot is processing information from the instruments outside of central vision, this viewing mode should hamper performance. Fox et al. compared performance of instructors and novices on a simulated IFR

flight task with and without preview. Performance of both novices and instructors was impaired when instrument preview was denied, and interestingly this impairment was larger for instructors than for novices. Two important conclusions can be drawn from these findings. First, it is apparent that pilots can process cockpit display information outside of central vision. Secondly, the extent to which they can do this depends on their flight expertise, pointing to a possible application in which measurements of the perceptual span are used as an indicator of expertise. Further research is encouraged to explore this potential.

A more recent study by Schaudt, Caulfield, & Dyre (2002) investigated the related issue of whether pilots could process information from cockpit information that was not fixated directly. They tested a method of presenting airspeed in pilots' peripheral vision and found that pilots were able to use this information, confirming that peripheral processing of cockpit information is possible. Research is needed to determine what kinds of information can be processed peripherally (e.g., position of instrument dials, colour information, text), and the extent to which this depends on expertise.

3.1.10 Event-related analysis

Another common approach to analyzing eye movements in the cockpit is to examine the changes that occur following an event (Alexander & Wickens, 2006; Björklund & Alfredson, 2006; Dehais, Causse, Régis, Menant, Labedan, Vachon, & Tremblay, 2012; Itoh et al., 1990; Matessa & Remington, 2005; Mumaw et al., 2001; Sarter et al., 2011; Schriver et al., 2008; Thomas & Wickens, 2004; van de Merwe et al., 2012; van Dijk et al., 2011). These cockpit events are typically inserted into a scenario during simulated flight, and are used as probes in order to measure the pilot's situational awareness. Situational awareness probe events are usually task-relevant, such as a system malfunction, a change in the aircraft's automation state, or the appearance of an object in nearby airspace. While it is instructive to require an overt behavioural response (e.g., verbal acknowledgment) to probe events, analysis of eye movement recordings can determine whether the pilot actually fixated the relevant event information within the cockpit or OTW. This is especially useful when it is inconvenient or impossible to collect an overt behavioural response from the pilot.

For example, Björklund and Alfredson (2006) used eye movement recordings to determine whether a pilot visually verified automation mode transitions. The authors found that the likelihood of gaze being directed to the cockpit indicator following the transition event differed between the Captain and First Officer (see also Mumaw et al., 2001; Sarter et al., 2011). Thomas and Wickens (2004) used eye tracking to find evidence that pilots had noticed the appearance of an object in nearby airspace. Dehais et al. (2012) used eye movements to determine whether a pilot's visually detected cockpit indications of a simulated landing gear failure. Itoh et al. (1990) monitored pilots' eye movements in a commercial airline simulator and found that following a system malfunction event, pilots showed an increase in total dwell duration on relevant cockpit display panels. Schriver et al. (2008) also examined eye movements following a system malfunction during simulated flight. They also showed increased dwell duration on relevant cockpit displays associated with 'noticing' the malfunction, and that there were differences in the latency to notice the malfunction between expert and novice pilots. A similar finding was observed by Matessa and Remington (2005) for malfunctions during simulated space shuttle launch operations. Van Dijk et al. (2011) found that eye movement entropy increased in the period following an instrument failure (see also van de Merwe et al., 2012). Hence eye

movements provide a variety of indicators of event-related information processing that can be used to assess situational awareness.

3.2 Models

Modeling efforts in this domain have sought to produce formal mathematical descriptions of the mechanisms that generate cockpit eye movements. While a detailed review of modeling of aircrew eye movements is beyond the scope of the present paper, there have been several modeling attempts in this domain that should be mentioned (for a review of early modeling in this domain see Jones, 1985). The earliest model of cockpit eye movement was developed by Senders (1964, 1966). This model assumed that instrument fixation probability was a zero-order Markov process, and attempted to predict these fixation probabilities based on the bandwidth of information presented in instruments. Subsequently, Carbonell and colleagues (Carbonell, 1966; Carbonell, Ward, & Senders, 1968; Carbonell, Senders, & Ward, 1969) devised a model to predict the proportion of time spent on each instrument. The model invokes a sampling process that determines a queue of instrument visits based on the risk of missing a critical observation from each instrument (though the model was not evaluated against real instrument dwell sequences and durations). In a later work, Jones (1985) constructed a model whose mechanics are based on the idea that the pilot seeks to minimize error from a desired state. The model was able to produce third-order transition matrices that matched those from actual pilots' eye movement data, and also produced mean dwell durations and entropy measures that showed agreement with real data.

Dick (1980) took a multivariate approach to modeling cockpit eye movement data during a simulated IFR landing scenario. Factor analysis was applied to a variety of eye movement measures including fixation duration, variability, transition probabilities, and blink rate, as well as simulator variables such as aircraft speed, deviation from glide slope, etc., across different segments of flight time. This analysis produced factors, which are collections of variables that co-vary. This approach is potentially very powerful in the context of cockpit eye movements, because it is likely that certain variables will be positively correlated with one another and exhibit redundancy (e.g., fixation frequency and total dwell duration), and others might be negatively correlated with one another (e.g., fixation frequency and mean dwell duration). Multivariate analyses such as factor analysis can extract the constellations of variables that change together, and also expose contrasts between variables, thereby illuminating the underlying relationships between multiple eye movement variables. For example, Dick (1980) derived factors that associated eye movement measures (e.g., mean dwell duration, variability, first-order transition values) with different levels of mental workload (e.g., manual control vs. autopilot) and different segments of flight. These factors could also be used to distinguish between individual pilots. He went on to describe a model in which pilots select 'mini-scan' monitoring patterns (described by aforementioned factors) based on the level of uncertainty about aircraft state.

Kramer et al. (1994) took a similar approach by applying discriminant analysis to multivariate eye movement data. In particular, they derived discriminant functions that used both eye movement variables (PDT, dwell frequency, 1st – 4th order Markov coefficients) as well as aircraft state variables (root-mean-squared error in velocity, altitude, etc.) in order to distinguish between student and instructor pilots. More recently the work by Hayashi (2004) applied hidden-Markov modeling to pilot cockpit transition matrices as a way to capture internal pilot control processes.

With this modeling approach, hidden-Markov states, which are conceptually similar to the factors emerging from factor analysis, code for pilot task (or strategy) and determine the subset of transitions that occur over a specific flight interval.

Doane and Sohn (2000) applied the ACT-R modeling framework in order to describe the cognitive activity of a pilot during simulated flight. This model (called ADAPT), constructs a representation of the pilot's knowledge about the aircraft state, and defines actions that result in updating of knowledge about the aircraft (e.g. look at cockpit indicator) changes to the aircraft state (e.g. control inputs). The model produced plans, or sequences of actions, including information acquisition activities, and Doane and Sohn demonstrated that these sequences can resemble the distribution of visual attention exhibited by pilots flying a simulated aircraft.

3.3 Other applications

Several applications of cockpit eye movement recordings are peripheral to the topics discussed previously, and will be covered in the following sections.

3.3.1 Usability evaluation

Several studies have used eye tracking to assist in the usability evaluation of a novel cockpit tool or instrumentation (Alexander & Wickens, 2006; Cote et al., 1985; Flemisch & Onken, 2000; Johnson et al., 2006; Hayashi, 2004; Oseguera-Lohr & Nadler, 2004; Schaudt et al., 2002; Schnell et al., 2004; Wickens et al., 2004). In an early application Cote et al. (1985) used eye tracking to compare flight performance using different possible navigation systems for a utility helicopter. Oseguera-Lohr and Nadler (2004) used eye movement measures (PDT and mean dwell duration) to assess the usability of a novel cockpit tool that assists with aircraft spacing during landing approach. Similarly, Wickens and colleagues (Alexander & Wickens, 2006; Johnson et al., 2006; Schnell et al., 2004; & Wickens et al., 2004) investigated changes in the distribution of pilots' visual attention in the context a novel Synthetic Vision System (SVS) that provides a heads-down synthetic view of the external environment. Measures of PDT provided insight into the extent to which subjects used the SVS versus the out-the-window view. A similar approach has been applied by Hayashi et al. (2009) to investigate the distribution of attention (both in PDT and pattern of dwells) within a fault management display in a next-generation spacecraft simulator. Two versions of a fault management display were compared and analysis of eye movements differentiated the displays in terms of eye movement pattern, and also the proportion of overall viewing time that was consumed by the display. In another study, Williams (2002) evaluated a novel 'highway-in-the-sky' synthetic flight display. Eye movements were recorded and PDT out-the-window was used as a measure of situational awareness. Hence eye movement recordings offer a way to determine the effects of new cockpit instrumentation on the distribution of visual attention within the cockpit.

3.3.2 Field-of-view requirements

In another application of eye movement recordings, Dixon, Rojas, Krueger, and Simcik (1990) sought to assess the field-of-view (FOV) requirements for flight simulators. To do this the authors compared eye movements of pilots who flew simulators that differed in the available out-the-window FOV (160˚ vs. 113˚ horizontal). Interestingly, they found that while the two FOV settings did not

produce different levels of performance, they did elicit different distributions of eye movements. This is an example of a situation where eye movement recordings were able to provide sensitivity in discriminating between experimental conditions where performance measures were not.

3.3.3 Qualitative feedback during training

Pilots' eye movements have also been used to provide qualitative feedback during training. Jones, Coates, and Kirby (1983) presented trainees with a video about the eye movements of expert pilots. While there was some indication that this influenced the trainees' self-reports about their own eye movement patterns, there was little evidence of an actual effect on trainees' flight performance. More recently, Wetzel, Anderson, and Barelka (1998) reported on eye movement recordings that had been integrated into a United States Air Force (USAF) basic F-16 simulator training regime. Specifically, instructors were able to monitor students' eye movements in real-time during flight performance and provide feedback regarding their instrument cross-check. Wetzel et al. surveyed the instructors and found that they regarded the eye tracking capability as a very useful component of the course. More recently, Carroll, Surpris, Strally, Archer, Hannigan, Hale, and Bennett (2013) suggested that eye movement recordings should be integrated into pilots' helmet-mounted display in order to provide after-action review (AAR) of pilot eye movement patterns during F-35 pilot training.

One of the primary difficulties with this instructor-feedback approach is that it requires the instructor to perceive the appropriate underlying information from the student's eye movement patterns while viewing them in real-time or else during an after-action review. Because eye movements are rapid, fixations short in duration, and dwell patterns potentially complex, it may be impossible for a human viewer to reliably extract the information upon which to give feedback (Carroll et al., 2013). Ideally, the information upon which this feedback might be based should be derived from quantitative eye movement measures such as those described in the previous sections. Nevertheless, an instructor might be able to give some useful qualitative feedback such as the pilot's eye movements following situational awareness probe events (e.g., verifying that the pilot looks at the automation state display following an aircraft automation mode change).

3.3.4 Aircraft inputs

Ineson, Durnell, Ebbage, Jarrett, Neary, and Reed (2004) explored the potential for eye tracking as an aircraft control input. The authors reasoned that under high G-forces, movements of the arm are more difficult than movements of the eye, and hence eye movements might be able to provide better means of aiming for pilots under these conditions. Subjects aimed at a target board under high G-load in a human centrifuge either under eye-controlled aiming or head-controlled aiming. Indeed, eye-controlled aiming was found to be unaffected by G-load whereas head-controlled aiming was. Eye-controlled aiming was also found to be faster than head-controlled aiming, though somewhat less accurate.

These results also point to a larger potential set of gaze-based human-computer interactions during aircraft training and operations. For example, an aircraft with integrated pilot eye movement monitoring might eventually be able to monitor pilot mental workload and situational awareness. This information could be recorded both for after-action review, but also for real-time pilot-aircraft interactions (e.g., an attention-aware instrument display that alerts the pilot when it has been ignored for too long).

4 Summary

More than half a century of research has demonstrated strong connections between eye movements in the cockpit and aviator cognitive state. This research has been conducted primarily using video-based eye trackers in simulators and in real aircraft, and for both civilian and military applications. Eye movements have been shown to provide indicators of expertise, mental workload, and situational awareness.

Perhaps one of the most robust findings from prior research is that expert aviators have shorter duration fixations and dwells than novices. Experts also make more frequent fixations, show a more varied and less predictable dwell pattern, and have a larger perceptual span than novices. There is also some evidence, though preliminary, that pupil diameter is sensitive to pilot expertise. These differences are likely to be due to the relative efficiency with which experts process cockpit information. Flight tasks or situations that result in increased mental workload are also associated with increased mean dwell duration, changes in the distribution of dwell duration over cockpit areas, and increased pupil diameter. Increased mental workload may also causes changes in the pattern of dwells (e.g., entropy), though this effect was not consistent across studies. Eye movements can also provide indications of the aviator's level of situation awareness. For example, several studies used eye movements to determine whether or not the pilot fixated relevant cockpit information during important flight events such as aircraft malfunctions or the appearance object in nearby airspace. Blinks may offer a measure of the pilot's level of fatigue. Hence for cockpit events in which no overt response is required (e.g., the operator merely needs to notice it), eye movements offer a unique contribution to depicting the operator's situational awareness; there may be no other way to capture pilots' "latent" information acquisition activity.

There are several other applications of eye movements in the cockpit that have been explored in prior research. Eye movements can aid in usability evaluation for novel cockpit instrumentation, tools and devices. Recordings of eye movements can be used as feedback during training, either in real-time or in an after-action review. This approach might seem intuitive however it does rely on the ability of the instructor to provide appropriate and helpful feedback. Furthermore, it might be difficult to recommend heuristics for such eye-movement feedback without conducting a quantitative analysis of some or more of the variables described earlier. Finally, eye movement monitoring can be used to provide gaze-based interactions, which is a class of promising yet largely unexplored cockpit applications.

Taken together, these findings strongly promote the potential utility of eye movement recordings for civilian and military flight. They could be applied during the training process to track and corroborate, together with performance measures, the management of workload and the development of expertise. Eye tracking could also be used to probe and verify pilot situational awareness, either during training or during actual flight performance. Given that eye tracking solutions are increasingly affordable and robust (see Annex B for a discussion), their application in these ways is feasible. Accordingly, a recent study by Air Force Research Laboratory has suggested that eye-tracking be integrated into a 5th-generation fighter aircraft (F-35) training program (Carroll et al., 2013). However, despite the large number of research studies that have examined eye movements in the cockpit, there are still many empirical and theoretical questions in this domain that warrant further research and these questions are outlined in the following section.

5 Future research

One important avenue for future research would be to achieve further empirical consensus and validation of eye movement measures in the cockpit. While some empirical eye movement measures showed robust relationships to aspects of the aviator's cognitive state, others produced a less clear pattern across studies. For example, the relationship between global pattern metrics such as dwell sequence entropy and entropy rate and operator variables such as expertise and mental workload were unclear across studies. Some of the variation in findings might be due to differences in flight task and mental workload manipulation across studies, and hence future studies might seek to examine dwell sequence pattern while carefully controlling these factors. In addition, measures such as pupil size, saccade length, and perceptual span during flight tasks have been investigated in only a few studies and would benefit from further empirical validation and replication. Other measures, such as microsaccade peak velocity, have not yet been explored in the cockpit and should be investigated.

Further work is also needed to characterize eye movements across different flight tasks and also across different individuals. Most eye movement measures were shown to be sensitive to flight task, indicating that flight context (e.g., task, role, cockpit design) is important to consider when interpreting patterns of cockpit eye movements. Aircrew roles and their corresponding task requirements are diversifying with the advent of new technology (e.g., pilot, navigator, sensor operator, UAV flight control) and accordingly it might be necessary to identify patterns of eye movement measures that correlate with workload, situational awareness, or expertise separately for different roles and tasks. In addition, there are strong individual differences in eye movements in the cockpit, particularly in dwell pattern. This tends to complicate analyses that seek common trends across subjects. One solution to this is to adopt a within-subject approach that examines eye movement measures relative to a pilot's own baseline. For research purposes this approach would yield more experimental power for comparisons across individuals, and for application design this approach would customizing the application to each individual's eye movement parameters. One promising analytic approach that emerged from prior work in this domain is multivariate analysis. In particular, factor analysis and other multivariate methods show great potential for a) reducing the large space of eye movement measures into tractable structures, b) capturing inter-measurement correlations, and c) including other external variables such as flight task and segment as well as information specific to the pilot, including expertise and performance variables. However, only a few studies in this domain have employed these techniques.

One of the most direct applications of eye movement recordings to aviator training is to have an instructor view the student's eye movements and provide feedback. As was discussed earlier, there are difficulties with this method that centre around the instructor's ability to extract important information by viewing the student's eye movement pattern. One avenue for future applied research would be to develop tools that would a) analyze and visualize students' eye movements (perhaps using some of the measures described earlier) and b) help highlight the connection of these measures to the task and performance (e.g., cockpit events, student flight performance). Such tools would make the instructor's job less qualitative and would potentially relieve the instructor of the difficult task of real-time viewing of eye movement records. Further research might seek to develop and validate quantitative guidelines, cut-offs, and criteria for eye movement measures that could be used as guidance for instructors.

Finally, there are many under-explored applications of eye movements to aircrew performance in the cockpit. As was covered earlier, one promising domain is gaze-based aircraft input. One possibility is hands-free or eye-assisted selection within aircraft multi-function displays. Moving beyond eye movements as control input, future research might develop aircraft instruments that are sensitive to, or that incorporate information about, the pilot's pattern of visual attention. Such attention-aware displays have been developed for a variety of computer-based applications (for a review see Toet, 2006; see also Rosch & Vogel-Walcutt, 2012), and many of these concepts could be applied in the aviation domain. Accordingly, in the future an attention-aware cockpit might emerge that is sensitive to its user's pattern of visual attention, providing better training (e.g., via expertise tracking; eye movements-based feedback), increased operational effectiveness (e.g., task- and operator-specific information display), and improved safety (e.g., sensitivity to operator mental workload and situational awareness).

This page intentionally left blank.

References

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics,* 36, 623-636.

Alexander, A. L., & Wickens, C. D. (2006). Integrated hazard displays: individual differences in visual scanning and pilot performance. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*.

Anders, G. (2001). Pilot's attention allocation during approach and landing – Eye- and head-tracking research in an A330 full flight simulator. Presentation at the 11th International Symposium on Aviation Psychology, Columbus, OH.

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine*, 68, 569-579.

Björklund, C., & Alfredson, J. (2006). Mode monitoring and call-outs: an eye tracking study of two-crew automated flight deck operations. *The International Journal of Aviation Psychology,* 16 (3), 263-275.

Brown, D. L., Vitense, H. S., Wetzel, P. A., & Anderson, G. M. (2002). Instrument scan strategies of F-117A Pilots. *Aviation, Space, and Environmental Medicine,* 73 (10), 1007-1013.

Burian, B. K., Pruchnicki, S., Rogers, J., Christopher, B., Williams, K., Silverman, E., Drechsler, G., Mead, A., Hackworth, C., & Runnels, B. (2013). Single pilot workload management in entry-level jets. NASA/TM-2013-216557.

Carbonell, J. R. (1966). A queueing model of many-instrument visual sampling. *IEEE Transactions on Human Factors in Electronics,* 7 (4), 157-164.

Carbonell, J. R., Senders, J. W., & Ward, J. L. (1969). Human sampling processes: A simulation-validation study. NASA CR-1258.

Carbonell, J. R., Ward, J. L., & Senders, J. W. (1968). A queueing model of visual sampling: Experimental validation. *IEEE Transactions on Man-Machine Systems*, MME-9, pp. 82-87.

Carroll, M., Surpris, G., Strally, S., Archer, M., Hannigan, F., Hale, K., Bennett, W. (2013). Enhancing HMD-based F-35 training through integration of eye tracking and electroencephalographic technology. In *D.D. Schmorrow and C.M. Fidopiastis (Eds.): AC/HCII 2013, LNAI 8027*, pp. 21-30. Springer-Verlag: Berlin Heidelberg, 2013.

Causse, M., Baracat, B., Pastor, J., & Dehais, F. (2011). Reward and uncertainty favor risky decision-making in pilots: evidence from cardiovascular and oculometric measurements. *Applied Psychophysiology and Biofeedback,* 36, 231-242.

Cheung, B., & Hofer, K. (2003). Eye tracking, point of gaze, and performance degradation during disorientation. *Aviation, Space, and Environmental Medicine,* 74 (1), 11-20.

Chuang, L. L., Nieuwenhuizen, F. M., & Bülthoff, H. H. (2013). A fixed-based flight simulator study: The interdependence of flight control performance and gaze efficiency. In: *Engineering Psychology and Cognitive Ergonomics: Applications and Services*, 10th International Conference EPCE 2013, Held as part of HCI International 2013, Springer, Berlin, Germany, 95-104.

Colvin, K., Dodhia, R. M., Belcher, S., Dismukes, K., Poly, C., & Obispo, S. L. Scanning for visual traffic: An eye tracking study. (2003). In *Proceedings of the 12th International Symposium of Aviation Psychology* (pp. 255-260). Dayton, OH: The Wright State University.

Colvin, K., Dodhia, R., & Dismukes, R. K. (2005). Is pilots' visual scanning adequate to avoid mid-air collisions? In *Proceedings of the 13th International Symposium on Aviation Psychology* (pp. 104-109). Oklahoma City.

Cote, D. O., Krueger, G. P., & Simmons, R. R. (1985). Helicopter copilot workload during nap-of-the-earth flight. *Aviation, Space, and Environmental Medicine,* 56 (2), 153-157.

Dahlstrom, N., & Nahlinder, S. (2009). Mental workload in aircraft and simulator during basic civil aviation training. *The International Journal of Aviation Psychology,* 19 (4), 309-325.

Dehais, F., Causse, M., Régis, N., Menant, E., Labedan, P., Vachon, F., & Tremblay, S. (2012). Missing critical auditory alarms in aeronautics: evidence for inattentional deafness? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50 (1)*, 1639-1643.

De Greef, T., Lafeber, H., van Oostendorp, H., & Lindenberg, J. (2009). Eye movements as indicators of mental workload to trigger adaptive automation. In *D.D. Schmorrow et al. (Eds.): Augmented Cognition, HCII 2009*, LNAI 5638, p. 219-228. Springer-Verlag Berlin Heidelberg, 2009.

Di Nocera, F., Camillia, M., and Terenzi, M. (2007). A random glance at the flight deck: pilots' scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making,* 1 (3), 271-285.

Di Stasi, L. L., McCamy, M. B., Catena, A., Macknik, S. L., Cañas, J. J., & Martinez-Conde, S. (2013). Microsaccade and drift dynamics reflect mental fatigue. *European Journal of Neuroscience*, 1-11. DOI:10.1111/ejn.12248.

Dick, A. O. (1980). Instrument scanning and controlling: using eye movement data to understand pilot behavior and strategies. NASA Contractor Report 3306; Contract NAS1-14793.

Diez, M., Boehm-Davis, D. A., Holt, R. W., Pinney, M. E., Hansberger, J. T., & Schoppek, W. (2001). *Tracking pilot interactions with flight management systems through eye movements*. Paper presented at the 11th International Symposium on Aviation Psychology, Columbus, Ohio.

Dinges, D. F., & Grace, R. (1998). PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance (*Rep. No. FHWA-MCRT-98-006*). Washington, DC: Federal Highway Administration, Office of Motor Carriers.

Dixon, K. W., Rojas, V. A., Krueger, G. M., & Simcik, L. (1990). Eye tracking device for the measurement of flight performance in simulators. *Final Report AFHRL-TP-89-18*, Air Force Human Resources Laboratory, Williams Air Force Base, Arizona, USA.

Doane, S. M., & Sohn, Y. W. (2000). ADAPT: A predictive cognitive model of user visual attention and action planning. *User Modeling and User-Adapted Interaction, 10*, 1-45.

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. 2nd Ed. London: Springer-Verlag.

Ellis, K. K. E. (2009). "Eye tracking metrics for workload estimation in flight deck operations." Thesis, University of Iowa, 2009. http://ir.uiowa.edu/etd/288.

Ellis, S. R. (1982). Contingency in visual scanning of cockpit traffic displays. In *Proceedings of the Human Factors Society 26th Annual Meeting*.

Ellis, S. R., & Stark, L. (1981). Pilot scanning patterns while viewing cockpit displays of traffic information. *Proceedings of the 17th Annual Conference on Manual Control, UCLA, June 16-18, 1981*, JPL Publication 81-95, pp 517-524.

Ephrath, A. R., Tole, J. R., Stephens, A. T., & Young, L. R. (1980). Instrument scan – Is it an indicator of the pilot's workload? *Proceedings of the Human Factors Society Annual Meeting,* 24, 257-258.

Findlay, J. M. (2009). Saccadic eye movement programming: sensory and attentional factors. *Psychological Research, 73*, 127-135.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences,* 22, 661-721.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. Aeronautical Engineering Review, 9 (2), 24-29.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1949). Eye fixations of aircraft pilots. *AF Technical Report No. 5967*, Wright-Patterson Air Force Base, Dayton, Ohio, USA.

Flemisch, O.F. and Onken, R., (2000). Detecting usability problems with eye tracking in airborne battle management support. In *Proceedings of the RTO HFM Symposium on "Usability of Information in Battle Management Operations,"* Oslo, Norway.

Fox, J., Merwin, D., Marsh, R., McConkie, G., & Kramer, A. (1996). Information extraction during instrument flight: An evaluation of the validity of the eye-mind hypothesis. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting.*

Gainer, C. A., & Obermayer, R. W. (1964). Pilot eye fixations while flying selected maneuvers using two instrument panels. *Human Factors,* 6, 485-501.

Glaholt, M. G., & Reingold, E. M. (2009). The time course of gaze bias in visual decision tasks. *Visual Cognition,* 17, 1228-1243.

Hankins, T. C., & Wilson, G. F. (1998). A comparison of heart rate, eye activity, EEG, and subjective measures of pilot mental workload during flight. *Aviation, Space, and Environmental Medicine,* 69 (4), 360-367.

Harris, R. L., Glover, B. J., & Spady, A. A. (1986). Analytical techniques of pilot scanning behavior and their application. *NASA Technical Paper,* 2525 19860018448, 39 p.

Hayashi, M. (2004). Hidden Markov models for analysis of pilot instrument scanning and attention switching. PH.D. Dissertation thesis; Massachusetts Institute of Technology.

Hayashi, M., Beutter, B., & McCann, R. S. (2005). Hidden Markov Model analysis for Space Shuttle crewmembers' scanning behavior. *IEEE International Conference on Systems, Man and Cybernetics*. Waikoloa, Hawaii, 2005.

Huemer, V. A., Hayashi, M., Renema, F., Elkins, F., McCandless, J. W., & McCann, R. S. (2005). Characterizing scan patterns in a spacecraft cockpit simulator: Expert vs. novice performance. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 83-87.

Ineson, J., Durnell, L., Ebbage, J. L., Jarrett, D. N., Neary, C., & Reed, M. A. (2004). Speed and accuracy of head- and eye-based aiming systems at high vertical acceleration. *Aviation, Space, and Environmental Medicine,* 75 (5), 420-428.

Itoh, Y., Hayashi, Y., Tsukui, I., & Saito, S. (1990). The ergonomic evaluation of eye movement and mental workload in aircraft pilots. *Ergonomics,* 33 (6), 719-732, DOI: 10.1080/00140139008927181.

Jessee, M. S. (2010). Ocular activity as a measure of mental and visual workload. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2010* 54: 1350.
DOI: 10.1177/154193121005401810.

Johnson, N., Wiegmann, D., & Wickens, C. (2006). Effects of advanced cockpit displays on general aviation pilots' decisions to continue visual flight rules flight into instrument meteorological conditions. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*.

Jones, D. H. (1985). An error-dependent model of instrument-scanning behavior in commercial airline pilots. NASA Contractor Report 3908. NASA, Langley.

Jones, D. H., Coates, G. D., & Kirby, R. H. (1983). The effectiveness of an oculometer training tape on pilot and co-pilot trainees in a commercial flight training program. NASA Contractor Report 3666. NASA: Langley.

Jones, D. G., & Endsley, M. R. (1996). Sources of situational awareness errors in aviation. *Aviation, Space, and Environmental Medicine,* 67 (6), 507-512.

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during VFR flight. *Proceedings of the 11th International Symposium on Aviation Psychology, Columbus, OH: The Ohio State University*.

Kantowitz, B. H. (1987). Mental workload. In P.A. Hancock (Ed.), *Human factors psychology* (pp. 81-131). Amsterdam: North-Holland.

Katoh, Z. (1997). Saccade amplitude as a discriminator of flight types. *Aviation, Space, and Environmental Medicine,* 68 (3), 205-208.

Kim, J., Palmisano, S. A., Ash, A., & Allison, R. S. (2010). Pilot gaze and glidescope control. *ACM Transactions on Applied Perception,* 7 (3), Article 18, 18 pp.

Kramer, A., Tham, M., Konrad, C., Wickens, C., Lintern, G., Marsh, R., Fox, J., & Merwin, D. (1994). Instrument scan and pilot expertise. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting – 1994.*

Krebs, M. J., Wingert, J. W., & Cunningham, T. (1977). Exploration of an oculometer-based model of pilot workload. NASA Contract Report 145153.

Latorella, K., Ellis, K. K. E., Lynn, B., Frasca, D., Burdette, D., Feigh, C. T., & Douglas, A. (2010). Dual Oculometer System for Aviation Crew Assessment. *Human Factors & Ergonomics Society Annual Meeting,* 2010.

Li, W.-C., Chiu, F.-C., Kuo, Y.-S., & Wu, K.-J. (2013). The investigation of visual attention and workload by experts and novices in the cockpit. In D. Harris (ed.), Engineering Psychology and Cognitive Ergonomics: Applications and Services. Springer: Berlin Heidelberg, pp. 167-176.

Li, W-C., Chiu, F-C., & Wu, K-J. (2012). The evaluation of pilots performance and mental workload by eye movement. *Proceedings of the 30th European Association for Aviation Psychology Conference*, Sardinia, Italy, 24-28 September, 2012.

Liu, Z., Yuan, X., Liu, W., & Rui, W. (2002). Characteristics of eye movement and cognition during simulated landing of aircraft. *Space Medicine & Medical Engineering,* 15 (5), 379-380.

Masson, G. S., & Stone, L. S. (2002). From following edges to pursuing objects. *Journal of Neurophysiology,* 88, 2869-2873. DOI: 10.1152/jn.00987.2001.

Matessa, M. M., & Remington, R. (2005). Eye movements in human performance modeling of space shuttle operations. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*.

McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics,* 17, 578-586.

McKinley, R. A., McIntire, L. K., Schmidt, R., Repperger, D. W., & Caldwell, J. A. (2011). Evaluation of eye metrics as a detector of fatigue. *Human Factors,* 53 (4), 403-414. DOI:10.1177/0018720811411297.

Morad, Y., Lemberg, H., Yofe, N., & Dagan, Y. (2000). Pupillography as an objective indicator of fatigue. *Current Eye Research,* 21, 535-542.

Morris, T. L., & Miller, J. C. (1996). Electrooculographic and performance indices of fatigue during simulated flight. *Aviation, Space and Environmental Medicine,* 42, 343-360.

Mumaw, R. J., Sarter, N. B., & Wickens, C. D. (2001). Analysis of pilots' monitoring and performance on an automated flight deck. In *Proceedings of the Eleventh International Symposium in Aviation Psychology* (p. 6). Columbus: Ohio State University.

Norris, J. R. (1998). Markov Chains. Cambridge University Press, Cambridge.

Oseguera-Lohr, R. M., & Nadler, E. D. (2004). Effects of an Approach Spacing Flight Deck Tool on Pilot Eyescan. NASA TM-2004-212987.

Ottati, W. L., Hickox, J. C., & Richter, J. (1999). Eye scan patterns of experienced and novice pilots during visual flight rules (VFR) navigation. *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*.

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences,* 88 (6): 2297–2301. doi:10.1073/pnas.88.6.2297

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology,* 32 (1), 3-25. DOI:10.1080/00335558008248231.

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). *Attention and the detection of signals. Journal of Experimental Psychology,* 109 (2), 160-174.

Previc, F. H., Lopez, N., Ercoline, W. R., Daluz, C. M., Workman, A. J., Evans, R. H., & Dillon, N. A. (2009). The effects of sleep deprivation on flight performance, instrument scanning, and physiological arousal in pilots. *The International Journal of Aviation Psychology,* 19 (4), 326-346.

Rayner, K. (2009). The thirty-fifth Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology,* 62, 1457-1506.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin,* 124 (3), 372-422.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 7, 65-81.

Rayner, K., Slattery, T. J., & Bélanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychological Bulletin and Review,* 17 (6), 834-839.

Reingold, E. M., & Charness, N. (2005). Perception in Chess: Evidence from Eye Movements. In *G. Underwood (Ed.). Cognitive processes in eye guidance* (pp. 325-354). Oxford: Oxford University Press.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive Psychology, 65 (2)*, 177-206. DOI: 10.1016/j.cogpsych.2012.03.001.

Robinski, M., & Stein, M. (2013). Tracing visual scanning techniques in training simulation for helicopter landing. *Journal of Eye Movement Research, 6 (2):3*, 1-17.

Rosch, J. L., & Vogel-Walcutt, J. J. (2012). A review of eye-tracking applications as tools for training. *Cognition, Technology, and Work, 15 (3)*, 313-127.

Russo, M., Thomas, M., Thorne, D., Sing, H., Redmond, D., Rowland, L., Johnson, D., Hall, S., Krichmar, J., & Balkin, T. (2003). Oculomotor impairment during chronic partial sleep deprivation. *Clinical Neurophysiology, 114*, 723-736.

Sarter, N. B., Mumaw, F. J., & Wickens, C. D. (2011). Pilots' monitoring strategies and performance on automated flight decks: an empirical study combining behavioral and eye-tracking data. *Human Factors, 49 (4)*, 347-357. DOI 10.1518/001872007X196685.

Schaudt, W. A., Caufield, K. J., & Dyre, B. P. (2002). Effects of a virtual air speed error indicator on guidance accuracy and eye movement control during simulated flight. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 46*: 1594. DOI: 10.1177/154193120204601714.

Schnell, T., Kwon, Y., Merchants, S., & Etherington, T. (2004). Improved Flight Technical Performance in Flight Decks Equipped With Synthetic Vision Information System Display. *The International Journal of Aviation Psychology, 14(1)*, 79–102.

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors, 50 (6)*, 864-878. DOI: 10.1518/001872008X374974.

Seeberger, J. J., & Wierwille, W. W. (1976). Estimating the amount of eye movement data required for panel design and instrument placement. *Human Factors, 18*, 281-292.

Senders, J. W. (1966). A re-analysis of pilot eye-movement data. *IEEE Transactions on Human Factors in Electronics, 7 (2)*, 103-106.

Senders, J. W. (1964). The human operator as a monitor and controller of multi-degree of freedom systems. *IRE Transactions on Human Factors in Electronics*, HFE-5, 2-5.

Shannon, C. E., & Weaver, W. (1949). *A Mathematical Model of Communication*. Urbana, IL: University of Illinois Press.

Spady, A. A. (1978). Airline pilot scan patterns during simulated ILS approaches. *NASA Technical Paper 1250*. NASA: Langley Research Center.

Spady, A. A., & Harris, R. L. (1981). How a pilot looks at altitude. *NASA Technical Memorandum 81967*. NASA: Langley Research Center.

Stern, J. A., & Bynum, J. A. (1970). Analysis of visual search activity in skilled and novice helicopter pilots. *Aerospace Medicine, 41*, 300-305.

Sullivan, J., Yang, J. H., Day, M., & Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation, Space, and Environmental Medicine, 82*, 871-878.

Svensson, E., Angelborg-Thanderz, M., Sjöberg, & Olsson, S. (1997). Information complexity – mental workload and performance in combat aircraft. *Ergonomics, 40 (3)*, 362-380.

Takeuchi, T., & de Valois, K. K. (2009). Visual motion mechanisms under low retinal illuminance revealed by motion reversal. *Vision Research, 9 (49)*, 801-809.

Thomas, E. L. (1964). The eye movements of a pilot during aircraft landing. *Aerospace Medicine, 34*, 424–426.

Thomas, L. C., & Wickens, C. D. (2004). Eye-tracking and individual differences in off-normal event detection when flying with a synthetic vision system display. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*.

Toet, A. (2006). Gaze directed displays as an enabling technology for attention aware systems. *Computers in Human Behaviour, 22*, 615-647.

Tole, J. R., Stephens, A. T., Vivaudou, M., Ephrath, A., & Young, L. R. (1983). Visual scanning behavior and pilot workload. *NASA-CR-3717*, NASA: Langley Research Center.

Tvaryana, A. P. (2004). Visual scan patterns during simulated control of an uninhabited aerial vehicle (UAV). *Aviation, Space, and Environmental Medicine, 75 (6)*, 531-538.

van de Merwe, K., Dijk, H. v., & Zon, R. (2012). Eye movements as an indicator of situational awareness in a flight simulator experiment. *The International Journal of Aviation Psychology, 22 (1)*, 78-95. DOI: 10.1080/10508414.2012.635129.

van Dijk, H., van de Merwe, K., & Zon, R. (2011). Situational awareness: studying relevant human factors tools. *The International Journal of Aviation Psychology, 21 (4)*, 343-356. DOI: 10.1080/10508414.2011.606747.

van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors, 43 (1)*, 111-121.

Veltman, J. A., & Gaillard, A.W.K. (1996). Physiological indices of workload in a simulated flight task. *Biological Psychology, 42*, 323-342.

Verdü, S. (2000). "Fifty years of Shannon theory." In S. Verdü and S. W. McLaughlin (Eds.) *Information Theory: 50 years of discovery*. IEEE Press. Pp. 13-34.

Vrzakova, H., & Bednarik, R. (2012). Hard lessons learned: Mobile eye-tracking in cockpits. *Gaze-In'12*, October 26, 2012, Santa Monica, California, USA.

Watson, A. B., & Yellot, J. I. (2012). A unified formula for light-adapted pupil size. *Journal of Vision, 12 (1)*, 1-16.

Weibel, N., Fouse, A., Emmenegger, C., Kimmich, S., & Hutchins, E. (2012). Let's look at the cockpit: exploring mobile eye-tracking for observational research on the flight deck. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, pp. 107-114, New York, NY, USA, 2012. ACM.

Wetzel, P. A., Anderson, G. M., & Barelka, B. A. (1998). Instructor use of eye position based feedback for pilot training. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting.*

Wickens, C. D., Alexander, A. L., Thomas, L. C., Horrey, W. J., Nunes, A., Hardy, T. J., & Zheng, X. S. (2004). Synthetic vision system display: The effects of clutter on performance and visual attention allocation. *NASA Technical Report AFHD-04-10/NASA(HPM)-04-1*, NASA Ames Research Center.

Wierwille, W. (1999). Historical perspective on slow eyelid closure: Whence PERCLOS? Paper presented at the *Ocular Measures of Driver Alertness Technical Conference*, Herndon, VA.

Wierwille, W., Rahimi, M., & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. *Human Factors, 27*, 489-502. DOI: 10.1177/001872088502700501.

Williams, K. W. (2002). Impact of aviation highway-in-the-sky displays on pilot situation awareness. *Human Factors, 44 (1)*, 18-27.

Wilson, G. F. (2002). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *International Journal of Aviation Psychology, 12 (1)*, 3-18.

Wilson, G. F., Caldwell, J. A., & Russell, C. A. (2007). Psychophysiological measures of fatigue effects on aviation related tasks of varying difficulty. *The International Journal of Aviation Psychology, 17:2*, 219-247. DOI: 10.1080/10508410701328839.

Wilson, G. F., & Fisher, F. (1991). The use of cardiac and eye blink measures to determine flight segment in F4 crews. *Aviation, Space, and Environmental Medicine, 62*, 959-962.

Wilson, G. F., Fullenkamp, P., & Davis, I. (1994). Evoked potential, cardiac, blink, and respiration measures of pilot workload in air-to-ground missions. *Aviation, Space, and Environmental Medicine, 65*, 100-105.

Yamada, E. (1969). Some structural features of the fovea centralis in the human retina. *Archives of Ophthalmology, 82*, 151-159.

Yang, J. H., Kennedy, Q., Sullivan, J, & Fricker, R. D. (2013). Pilot performance: Assessing how scan patterns& navigational assessments vary by flight expertise. *Aviation, Space and Environmental Medicine, 84 (2)*, 116-124.

Yarbus, A. L. (1967). *Eye movements and vision* (B. Haigh, Trans.). New York: Plenum Press.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115 (4)*, 787-835.

This page intentionally left blank.

Annex A Summary table

Table A.1: Chronological summary of the empirical studies that were considered in this review.

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Fitts, Jones, & Milton (1949)	35mm video camera	Cockpit-mounted	C-45 aircraft	Landing approach	Pilot experience	Fixation duration, fixation frequency, fixation pattern
Fitts, Jones, & Milton (1950)	35mm video camera	Cockpit-mounted	C-45 aircraft	Landing approach	Pilot experience	Fixation duration, fixation frequency, fixation pattern
Gainer & Obermayer (1964)	PAR Head Mount Visual Recording Camera	Head-mounted	Simulated YF-102	Variety of flight maneuvers	Instrument configuration	Fixation duration, fixation frequency, fixation pattern
Senders (1966)	Analyzed data from Fitts et al. (1949)	-	-	-	Modeling	Fixation duration, fixation frequency, fixation pattern
Stern & Bynum (1970)	Electro-oculogram (EEG)	-	UH-1D Helicopter	Cross-country flight	Pilot skill	Blink rate, fixation pattern
Krebs, Wingert, & Cunningham (1977)	Honeywell Mark IIA Oculometer	Cockpit-mounted	737 simulator	IFR approach and landing	Flight task difficulty, workload	Fixation duration, proportional dwell time, pupil diameter, blink rate, saccade amplitude
Spady (1978)	Honeywell Mark IIA Oculometer	Cockpit-mounted	737 simulator	Landing approach	Mental Workload, level of automation, task difficulty (turbulence)	Dwell time, fixation pattern

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Dick (1980)	Honeywell Oculometer	Cockpit-mounted	737 simulator	IFR landing approach	Mental workload, level of automation, task difficulty (turbulence), modeling	Fixation duration, proportional dwell time, dwell time variability, fixation pattern
Ephrath et al. (1980)	Honeywell Oculometer	Cockpit-mounted	Simulator	Landing approach	Mental workload, task complexity	Dwell pattern
Spady & Harris (1981)	Custom Oculometer	Cockpit-mounted	737 simulator	Landing approach	Instrument design	Proportional dwell time, dwell duration, dwell frequency
Tole et al. (1983)	Modified Honeywell Oculometer	Cockpit-mounted	ATC-510 procedures training simulator	IFR flight straight and level	Mental workload	Dwell duration, dwell pattern
Cote et al. (1985)	NAC Eye Mark Recorder	Head-mounted	JUH-IH helicopter	Nap-of-the-earth flight	Navigation systems compared	Fixation frequency
Jones (1985)	Modified Honeywell Oculometer	Cockpit-mounted	737 simulator	Steep turn maneuver	Modeling	Dwell duration, dwell frequency, dwell pattern
Wierwille et al. (1985)	Video camera	Cockpit-mounted	GAT-1B training simulator	Straight and level flight	Mental workload	Blink frequency, pupil diameter
Harris et al. (1986)	Review paper	-	-	-	Mental workload, expertise	Dwell duration, dwell frequency, dwell pattern
Dixon et al. (1990)	ASL 210	Head-mounted	C-130 trainer; F15/F16 simulator	Basic contact maneuvers	Field of view requirements for simulator	Proportional dwell time, dwell frequency

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Itoh et al. (1990)	NAC EMR-V Eye Mark Recorder	Head-mounted	767, 737-300 simulators	Routine flight with equipment malfunction	Mental workload	Fixation duration, fixation pattern, event-related analysis.
Wilson & Fisher (1991)	Electro-oculogram (EOG)		F4 Aircraft	Basic flight maneuvers	Flight segment classification	Blink rate, blink duration and other related variables.
Kramer et al. (1994)	ASL 4000H	Head-mounted	Custom flight simulator	Straight flight, turn maneuvers	Expertise	Fixation duration, fixation frequency, fixation pattern
Wilson et al. (1994)	Electro-oculogram (EOG)	-	F4 aircraft	Simple flight	Mental Workload	Blink rate, blink duration
Fox et al. (1996)	SRI Purkinje Gen5	Desktop-mounted (with bite bar)	Desktop flight simulator	Straight flight, turn maneuvers	Expertise, perceptual span	Fixation duration, fixation frequency, fixation pattern
Veltman & Gaillard (1996)	Electro-oculogram (EOG)	-	F-18 simulator	Straight flight, turn maneuvers	Mental workload	Blink duration, blink frequency
Bellenkes et al. (1997)	ASL 4000	Head-mounted	Beach Sport (Sundowner) simulator	Straight flight, turn maneuvers	Expertise, flight task	Dwell frequency, dwell duration
Katoh (1997)	Electro-oculogram	-	Custom flight simulator	Take-off, landing, straight flight, maneuvers	Flight task/mental workload	Saccade amplitude, dwell duration

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Svensson et al. (1997)	Video camera	Cockpit-mounted	JA37 flight simulator	Fly course, threat information handling	Mental workload	Fixation duration
Hankins & Wilson (1998)	Electro-oculogram (EOG)	-	Piper Arrow aircraft	Takeoff, approach, landing	Flight task/mental workload	Blink rate
Wetzel et al. (1998)	El Mar Inc. Vision 2000	Head-mounted	F-16 simulator	Basic flight	Real-time eye movement feedback during training; cross-check evaluation	Qualitative analysis
Ottati et al. (1999)	ALS E 5000	Head-mounted	Desktop flight simulator	Fly course	Expertise	Fixation duration, fixation frequency
Flemisch & Onken (2000)	Custom video-based eye tracker (caSBAro)	Head-mounted	Custom transport aircraft flight simulator	Fly course	Instrument display arrangement	Proportional dwell time
Anders (2001)	SMI iView-HED+HT	Head-mounted	A330 simulator	Approach, landing	Flight task, individual differences	Proportional dwell time
Diez et al. (2001)	ASL 504	Head-mounted	Aerowinx PS1 desktop simulator	Takeoff, Descent, Landing approach	Situational awareness, individual differences	Fixation duration, proportional dwell time
Kasarskis et al (2001)	ASL 5000	Head-mounted	‘Fly!’ desktop simulator	VFR landing approach	Expertise	Fixation frequency, dwell duration
Mumaw et al. (2001)	Not reported	-	747-400 simulator	Basic flight maneuvers	System automation	Proportional dwell time

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Brown et al. (2002)	El Mar Inc. Vision 2000	Head-mounted	F117A simulator	Straight flight, turn maneuvers	Scan strategy, individual differences	Fixation frequency, proportional dwell time
Liu, Yuan, Liu, & Rui (2002)	Xi'an Electrical Eye Movement Measure System (EMMS)	Not reported	Desktop F-15 simulation	Landing	Landing stages	Saccade velocity
Schaudt et al. (2002)	ASL 5000	Head-mounted	Custom flight simulation	Fly course	Effectiveness of virtual speed indicator (usability); peripheral processing	Proportional dwell time, fixation frequency
Williams (2002)	El Mar Inc. Vision 2000	Head-mounted	Custom flight simulator	Fly course	Effectiveness of route guidance displays (usability)	Proportional dwell time
Wilson (2002)	Electro-oculogram	-	Piper Arrow aircraft	VFR/IFR takeoff, flight, approach	Mental workload/flight task	Blink rate
Cheung & Hofer (2003)	Elmar Inc. Vision 2000	Head-mounted	Custom flight simulator	Fly course	Disorientation	Saccade frequency, pupil diameter
Colvin et al. (2003)	ISCAN Inc. Light Of Sight	Head-mounted	AST Hawk 201 flight training device	VFR flight	Mental workload	Proportional dwell time
Hayashi (2004)	ISCAN Inc. eye tracker	Head-mounted	Microsoft Flight Simulator (simulating 757-200)	IFR landing approach	Evaluate display concept (usability), expertise, modeling	Proportional dwell time, dwell pattern
Ineson et al. (2004)	Custom eye tracker	Head-mounted	G-simulation in centrifuge	Simple aiming task	Eye pointing vs. hand pointing	No eye movement measures reported

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Oseguera-Lohr & Nadler (2004)	ISCAN AA-ETL-500	Head-mounted	B-757 Simulator	Fly course	Evaluate navigation tool (usability), individual differences	Proportional dwell time, fixation pattern
Schnell et al. (2004)	ISCAN ETL 500	Head-mounted	Custom flight simulator	Landing approach	Evaluate display concept (usability), situational awareness, mental workload	Fixation duration, fixation frequency, saccade amplitude, fixation pattern
Tvaryana (2004)	El Mar Inc. Vision 2000	Head-mounted	RQ-1 Predator simulator	Fly course	Task difficulty/mental workload	Dwell duration, dwell frequency
Wickens et al. (2004)	ASL eye tracker (unspecified model)	Not reported	Frasca simulator	VFR and IFR Landing approach	Evaluate display concept (usability), situational awareness, individual differences	Dwell duration, dwell frequency, event-related analysis
Colvin et al. (2005)	ISCAN Line Of Sight	Head-mounted	AST Hawk 201 flight training device	VFR fly course	Task difficulty/mental workload	Proportional dwell time
Hayashi et al. (2005)	ISCAN ETL-500	Head-mounted	Space Shuttle part-task simulator	Launch simulation	Flight segment, individual differences	Dwell pattern
Huemer et al. (2005)	ISCAN ETL-500	Head-mounted	Space Shuttle part-task simulator	Identify system malfunction	Expertise	Proportional dwell time
Matessa & Remington (2005)	ISCAN ETL-500	Head-mounted	Space shuttle simulator	Identify system malfunction	Expertise	Dwell duration, fixation frequency, event-related analysis

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Alexander & Wickens (2006)	ASL 5000	Head-mounted	Frasca simulator	Fly course, detect events	Evaluate display concept (usability), individual differences	Proportional dwell time, event-related analysis
Björklund & Alfredson (2006)	GazeTracker	Head-mounted	Boeing 737NG simulator	Fly course, detect failure event	Automation, mental workload, individual differences	Dwell frequency, event-related analysis
Johnson et al. (2006)	Not reported	-	Custom desktop flight simulator	Fly course	Evaluate display concept (usability), pilot decision making, individual differences	Proportional dwell time
Di Nocera et al. (2007)	Tobii ET17	Desktop-mounted	Desktop flight simulator (Microsoft Flight Sim. 2004)	Climb, Cruise, Descend	Task difficulty/mental workload	Fixation pattern
Wilson et al. (2007)	SR Research Eyelink II	Head-mounted	Desktop UAV part-task simulator	Cruise, Monitor target area, target evaluation	Fatigue, task difficulty/mental workload	Pupil size
Schriver et al. (2008)	ASL 501	Head-mounted	Frasca 142 simulator	Fly course, detect system failure	Expertise	Proportional dwell time, event-related analysis

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Dahlstrom & Nahlinder (2009)	Electro-oculogram	-	Piper Arrow 31 Navajo aircraft; Flight Navigation Procedure Trainer	Takeoff, level flight, IFR approach	Expertise, workload, real aircraft vs. simulator	Eye movement energy
Ellis (2009)	SmartEye 3-camera system	Cockpit-mounted	737-800 simulator	Landing approach	Workload, automation	Fixation duration
Hayashi et al. (2009)	ISCAN ETL-500	Head-mounted	Orion spacecraft simulator	Spacecraft takeoff and ascent	Fault monitoring, situational awareness	Proportional dwell time, dwell duration, dwell frequency, dwell pattern.
Previc et al. (2009)	ASL Eye-Trac 6000	Head-mounted	T-6 aircraft simulator	Takeoff, climb, turn, descent, landing	Fatigue	Fixation duration, proportional dwell time, saccade length, pupil diameter, blink rate
Jessee (2010)	ASL Eye-Trac 6	Head-mounted	UH-60M Blackhawk simulator	Hovering flight, actions on contact	Task difficulty/mental workload	Fixation duration, fixation duration variability, saccade amplitude, blink interval
Kim et al. (2010)	Point Grey Research Inc. Firefly-MV Firewire camera	Desktop-mounted	Desktop Cessna simulator	Landing approach	Expertise	Fixation duration

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Latorella, Ellis, Lynn, Frasca, Burdette, Feigh, & Douglas (2011)	SmartEye Pro 5.5	Cockpit-mounted	Custom flight simulator	Basic flight	Evaluate eye tracking system performance	Eye track quality
Causse et al. (2011)	Pertech eye tracker	Head-mounted	Simple desktop ILS simulation	IFR landing approach	Landing decision making	Dwell duration
McKinley et al. (2011)	EyeCom Inc. EC6	Cockpit-mounted	UAV simulation	Landing	Fatigue	Percent eye closure, Approximate Entropy
Sarter et al. (2011)	ASL 4000	Head-mounted	B-747-400 simulator	Takeoff, flight, landing	Automation	Proportional dwell time
Sullivan et al. (2011)	Seeing Machines Inc. FaceLAB	Cockpit-mounted	Desktop helicopter simulator	Fly course	Expertise, Navigation	Dwell duration, Dwell pattern
van Dijk et al. (2011)	ASL 6000	Head-mounted	Custom flight simulator	Basic flight, diagnose malfunction	Situational awareness	Proportional dwell time, dwell pattern, event-related analysis
Dehais et al. (2012)	Pertech	Head-mounted	Custom flight simulator	Landing	Situational awareness	Event-related analysis, Saccade frequency
Li et al. (2012)	ASL Mobile Eye	Head-mounted	Custom flight simulator	Fly course (IFR)	Mental workload, situational awareness	Total dwell time
van de Merwe et al. (2012)	ASL 6000	Head-mounted	Custom flight simulator	Fly course	Situational awareness	Fixation duration, dwell duration, fixation pattern, fixation frequency, event-related analysis
Vrzakova & Bednarik (2012)	Ergoneers Ltd. Dikablis	Head-mounted	Airbus Jet simulator	Not reported	Evaluate eye tracking system	Calibration/Track quality (qualitative)

Citation	Eye tracker	Eye tracker Mount	Flight setting	Flight task	Factors considered	Eye movement measures
Weibel, Fouse, Emmenegger, Kimmich, & Hutchins (2012)	Tobii Glasses System	Head-mounted	Airliner simulator	Not reported	Evaluate eye tracking system/ analysis package	None reported
Burian, Pruchnicki, Rogers, Christopher, Williams, Silverman, Drechsler, Mead, Hackworth, & Runnels (2013)	Seeing Machines Inc. FaceLAB v5	Cockpit-mounted	ELJ Level 5 flight training device	Takeoff, flight, landing	Pilot error, mental workload	None reported
Chaung et al. (2013)	Seeing Machines Inc. FaceLAB	Cockpit-mounted	Desktop flight simulator	Fly course	Task difficulty, individual differences	Dwell frequency, dwell pattern
Li et al. (2013)	ASL Mobile Eye	Head-mounted	Dynamic High Fidelity Trainer	Air-to-air combat	Mental workload, expertise	Dwell duration, proportional dwell time
Robinski & Stein (2013)	Ergoneers Ltd. Dikablis	Head-mounted	Eurocopter EC135 simulator	Takeoff, landing	Task difficulty, expertise	Proportion long gazes
Yang et al. (2013)	Seeing Machines Inc. FaceLAB 4.6	Cockpit-mounted	Desktop simulator	Fly course	Task difficulty, expertise	Dwell duration, dwell pattern

Annex B Eye-tracking technology primer

An eye tracker monitors a viewer's point of gaze in real time. While a detailed review of the history of eye tracking technology is outside of the scope of the present paper (see Duchowski, 2007), the present Annex serves as a primer on the capabilities of current eye tracking technology.

Most modern eye-tracking systems use a video camera to record images of a subject's eye (or eyes) in real time and estimate their gaze position in the world. This technique is possible because the appearance of the eye changes as the viewer looks at different points in the visual field. In particular, the shape of the pupil changes from being circular when staring straight at the camera to becoming more elliptical as the viewer gazes more eccentrically relative to the camera's line of sight. In addition, most video-based eye tracking systems illuminate the eye with an infrared light source. This provides steady illumination of the eye, making a more reliable image, but also creates reflections on the surface of the eye (the cornea) known as Purkinje reflections. The appearance of these reflections also varies with the eye position (or gaze direction) and hence can be used to predict gaze location. Most modern eye trackers use information about both the pupil and corneal reflection to track gaze location (the Pupil-CR method).

In order to track gaze position in the real world (i.e., identify what region of space the viewer is looking at), the eye tracker must translate changes in image characteristics of the eye (Pupil-CR appearance) to real world gaze locations. This is achieved through a calibration process. Basically, calibration involves having the viewer look at known points in the visual field (usually displayed on a computer screen), recording the eye image characteristics for each point, and then solving mathematical equations to produce a model that predicts gaze position within the plane of calibration based on the eye image characteristics. All video based eye trackers require some form of calibration, though some systems use defaults, or automatic calibration methods that simplify the process. Calibration is typically a critical factor in determining the system's accuracy in estimating the viewer's gaze position.

Because the appearance of the pupil and the corneal reflectance will also change with head movements, head movements are a confounding variable for eye trackers. Solving this issue has led to one of the major design distinctions in eye tracking technology: head (helmet, or glasses) mount vs. desk (or cockpit) mount. Head-mounted eye trackers eliminate the problem of head movement by mounting the eye-tracking camera on the head, so the relative image of the eye remains approximately constant. This is a good solution save for a few drawbacks: 1) the head-mounted camera can be bulky and obtrusive, and 2) the camera may still move if the head-mount shifts during head movements. Also, for head-mounted eye trackers, the position of the head must also be monitored in order to translate the recorded eye information into real-world gaze positions (e.g., calibration). For desk-mounted eye trackers, head movements are often mitigated by using a chinrest/forehead support or a bite-bar. This method is effective at reducing head movements but can also be quite obtrusive and/or uncomfortable. More recently, desk (or cockpit) mounted eye tracking solutions have emerged that track the eye as well as the head, and thereby compensate for head motion in their solution for gaze position. This can involve tracking features of the face, or overtly tracking head position with a separate head-tracker.

Perhaps the most important differences between eye trackers stem from the quality of the eye tracking data that they provide. Eye trackers vary in accuracy and precision of gaze position measurement, as well as temporal frequency. Accuracy is the average error in predicting gaze position; for high-fidelity eye tracking systems this is < 0.5 degrees of visual angle. This means that the system is able to estimate the target of the viewer's gaze within 0.5 degrees. Precision corresponds to the smallest change in gaze position that is detectable by the system, and is often < 0.1 degrees of visual angle. This is important for detecting movements of the eye (e.g., saccades). Temporal frequency is the rate at which the eye tracker samples gaze position, and ranges from relatively slow (25 Hz) to extremely fast (2000 Hz). Faster sampling rate allows the eye tracker to detect very short duration eye movement events (e.g., microsaccades). Another critical variable for applying eye tracking systems is the range over which they can track head and eye position. This is particularly important for desk- or cockpit-mounted systems; it corresponds to how far the subject can move their head, or turn their head, and the area over which the subject can fixate within the visual field. For a cockpit eye tracking system, a wide tracking range is desirable. Note that these factors are not necessarily independent. For example, increased range often comes at the expense of accuracy, and temporal sampling rate and accuracy may show an empirical tradeoff. Eye tracking systems can be sensitive to ambient light; some eye trackers must be used in a setting of relatively controlled illumination and will perform poorly when used in variable illumination (e.g., outdoors). Eye tracking systems may also be more or less tolerant to corrective lenses (particularly glasses, which can be reflective). Hence when evaluating eye tracking solutions for a simulator or aircraft cockpit all of the aforementioned factors should be considered.

As can be seen in the summary table in Annex A, studies on eye movements in the cockpit have used both head-mounted and cockpit-mounted systems. Both types of systems have drawbacks for the cockpit setting that must be considered: head-mounted systems might interfere with other head-mounted devices worn during flight (e.g., a helmet; night vision goggles mounted on the helmet), while cockpit-mounted systems might interfere with the information displayed in the cockpit. Beyond these concerns, it is important to determine whether the system can provide the measurements that are needed for the research experiment or application. For example, to measure percent dwell time on individual instrument panels, the eye tracker must be able to track gaze position with an appropriate accuracy. If, when sitting in the cockpit, an instrument panel occupies one degree of the pilot's visual angle, an eye tracker must have accuracy on the order of one degree or better to determine whether the pilot is looking at that instrument versus an adjacent instrument. If microsaccades are of interest, the eye tracker will need to have high accuracy, precision, and a high temporal sampling rate in order to detect these small, rapid movements. This paper does not evaluate specific models of eye trackers used in the cockpit, thought the specific model of eye tracker used in each of the cited empirical studies can be found in the table in Annex A. Several studies have tested specific eye tracking systems in the cockpit and reported on their applicability (SmartEye, Latorella et al., 2011; Dikablis, Vrzakova & Bednarik, 2012; Tobii Glasses, Weibel et al., 2012; see also Carroll et al., 2013).

Finally, it should be noted that there are other methods than the video-based method for eye tracking. One method that has been used with some frequency in the context of cockpit eye movement recording is the electro-oculogram (EOG). This method involves recording oculomotor muscle activation, using electrodes that are placed on the face and head. This method can accurately detect the onset of saccades in the vertical and horizontal planes. However, there are several drawbacks to this method. Because it involves recording muscle activity, the data are

noisy, and the requirement to wear electrodes on the head can also create conflicts with other devices. Electrodes might also impede upon comfort, and might require a technician to apply/monitor. Furthermore, the exact gaze position in the cockpit is difficult to estimate from EOG, and hence in prior research this technique has primarily been used to detect the saccades and blinks.

List of symbols/abbreviations/acronyms/initialisms

AAR	After Action Review
ApEn	Approximate Entropy
D Air Pers Strat	Director Air Personnel Strategy
DRDC	Defence Research and Development Canada
EOG	Electro-oculogram
FOV	Field Of View
IFR	Instrument Flight Rules
NASA	National Aeronautics and Space Administration
NNI	Nearest Neighbor Index
OTW	Out The Window
PDT	Percent Dwell Time
RCAF	Royal Canadian Air Force
SVS	Synthetic Vision System
UAV	Unmanned Aerial Vehicle
USAF	United States Air Force
VFR	Visual Flight Rules

DOCUMENT CONTROL DATA

(Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g., Centre sponsoring a contractor's report, or tasking agency, are entered in Section 8.) DRDC – Toronto Research Centre Defence Research and Development Canada 1133 Sheppard Avenue West Toronto, Ontario M3K 2C9 Canada	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) UNCLASSIFIED 2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC DECEMBER 2012

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Eye tracking in the cockpit: a review of the relationships between eye movements and the aviator's cognitive state.

4. AUTHORS (last name, followed by initials – ranks, titles, etc., not to be used)

Glaholt, M. G.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
December 2014	58	128

7. DESCRIPTIVE NOTES (The category of the document, e.g., technical report, technical note or memorandum. If appropriate, enter the type of report, e.g., interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Scientific Report

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

DRDC – Toronto Research Centre
Defence Research and Development Canada
1133 Sheppard Avenue West
Toronto, Ontario M3K 2C9
Canada

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
03rm	
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)	**10b. OTHER DOCUMENT NO(s).** (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
DRDC-RDDC-2014-R153	

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

The eye movements of aircrew during flight have been a topic of interest to military and civilian researchers for over 60 years. Studies in flight simulators and real aircraft have used eye movements as a window onto operators' processing of information from cockpit instruments and displays. This body of research has demonstrated strong links between eye movements and pilot expertise, workload, and situational awareness. Other applications of eye tracking in the cockpit have also been proposed, such as usability analysis, instructor-feedback, and aircraft input. The present report reviews this body of research to date and provides recommendations for future research and applications.

Les mouvements oculaires des équipages en vol suscitent l'intérêt des chercheurs militaires et civils depuis plus de 60 ans. Au cours d'études menées en simulateurs de vol et à bord d'aéronefs, ils ont servi à analyser la façon dont les pilotes traitent l'information fournie par les instruments et les dispositifs d'affichage de leur habitacle. Ces travaux ont montré qu'il existe des liens étroits entre le mouvement des yeux des pilotes et leur savoir-faire, leur charge de travail et leur connaissance de la situation. D'autres applications de suivi des mouvements oculaires ont été proposées, celles-ci portant notamment sur l'analyse de la facilité d'utilisation, sur la rétroaction d'instructeurs et sur les données fournies par les aéronefs. Le présent rapport traite de la recherche effectuée jusqu'à maintenant à ce chapitre et présente des recommandations en matière de recherches et d'applications futures.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g., Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

eye movements, eye tracking, aircrew, training, expertise, workload